REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE Feb. 18, 1999	3. REPORT TYPE AND DATES COVERED FINAL REPORT 02/05/98-01/05/99

4. TITLE AND SUBTITLE

NOVEL BIOPOLYMERS BASED ON AN EXPANDED GENETIC ALPHABET

5. FUNDING NUMBERS

N00014-98-1-0446

6. AUTHOR(S)

STEVEN A. BENNER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Florida
Department of Chemistry
P.O. Box 117200/440 Leigh Hall
Gainesville, FL 32611-7200

8. PERFORMING ORGANIZATION REPORT NUMBER

1666052-12

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

19990226020

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Problems encountered by the Navy can frequently be solved if only a specific receptor or catalyst could be discovered or developed for a particular ligand or a receptor. In these cases, the principal obstacle to solving the problem is often the unavailability of the receptor or catalyst in Nature, and the limits of chemical theory that prevent design. The equipment purchased under this grant was used to develop a method that promises to offer the Navy receptors and catalysts "on demand", using a technology known as *in vitro* selection build on an expanded genetic information system, whose chemistry and enzymology was developed in the Benner laboratories over the past few years.

14. SUBJECT TERMS

DTIC QUALITY INSPECTED 1

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

Grant #: N0014-98-1-0046

PRINCIPAL INVESTIGATOR: Steven A. Benner

INSTITUTION: University of Florida

GRANT TITLE: Novel Biopolymers Based on an Expanded Genetic Alphabet

AWARD PERIOD: 02/05/98-01/05/99

OBJECTIVE: To purchase equipment to support research funded under an ONR grant. This research is developing tools to select new catalysts, receptors and ligands based on an expanded genetic information system.

APPROACH:
   The DNA alphabet is not limited to the four standard nucleotides known in natural DNA. Rather, twelve nucleobases forming six base pairs joined by mutually exclusive hydrogen bonding patterns are possible within the geometry of the Watson-Crick base pair. The PI and his group have developed the chemistry and enzymology of these nucleobases, which are now known as An Expanded Genetic Information System (AEGIS). AEGIS is a new, "rule based" molecular recognition system that carries much of the functionality of proteins, can be prepared in combinatorial form, and can be copied and amplified much like nucleic acids. Figure 1 shows some of the **standard** and **non-standard** base pairs.
   One of the "Holy Grails" in the molecular sciences is a method to generate receptors, ligands, and catalysts "on demand", without the need for design. A variety of combinatorial technologies has emerged that search for these molecules among a collection of molecules of random structure. For example, a combinatorial technique known as *in vitro* selection (IVS) has been developed by Szostak, Ellington, Gold, Joyce, Cech and others to obtain DNA and RNA receptors, ligands (often called aptamers), and catalysts (often called ribozymes) on demand (Figure 2). The approach begins with a library of oligonucleotides of random sequence, typically 10^{14} individual DNA molecules. The library is passed through an affinity column bearing the "target" receptor. Oligonucleotides from the library that bind to the receptor are retained on the column, to be later eluted and amplified using the polymerase chain reaction (PCR). The amplification is coupled with mutation, to generate a new library of oligonucleotides, related by descent to DNA molecules that had some affinity to the receptor. Some of the mutants bind the receptor better than their parents; some bind worse. Repeating the selection extracts the former preferentially, and the cycle is repeated to yield aptamers that bind to the target receptor.
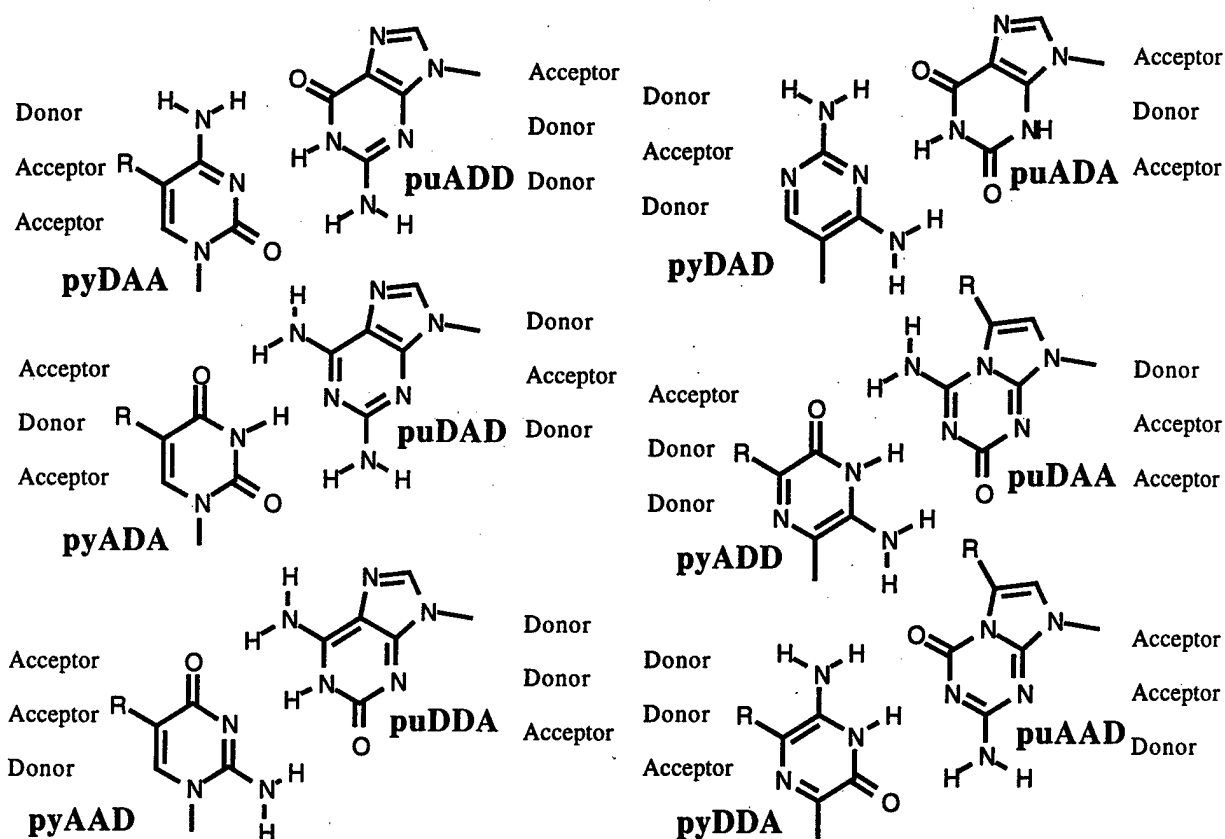
Figure 1. Twelve possible nucleobases in a DNA- or RNA-based "alphabet" that can form specific base pairs within the constraints of the Watson-Crick base pair geometry. Pyrimidine base analogs are designated by "py", purine by "pu". The upper case letters following the designation indicate the hydrogen bonding pattern of acceptor (A) and donor (D) groups. Thus, the standard nucleobase cytosine is pyDAA, guanosine is puADD (for example). R indicates appended functional groups.
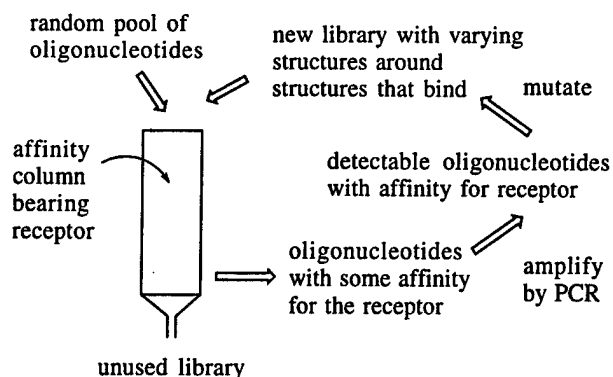

Figure 2. A schematic diagram describing *in vitro* selection for a DNA molecule that binds to an immobilized receptor.

If a small molecule ligand is attached to the column, then the procedure can yield DNA and/or RNA receptors that bind

the ligand. Likewise, "ribozymes" and "deoxyribozymes" that catalyze reactions can be obtained by IVS if, through its catalytic power, an oligonucleotide that catalyzes a reaction is separated from oligonucleotides that do not.

DNA- and RNA-based ligands and receptors have been raised against a variety of targets, including nucleic acids, proteins, peptides and small organic molecules. However, a decade of experimental work has shown that the catalytic power of ribozymes and the affinity of aptamers is frequently disappointing, especially for reactions and ligands that are not themselves nucleic acids.

The limitations of deoxyribozymes generated by IVS experiments are due in large part to the absence of chemical functionality on DNA and RNA, especially when compared with the functionality found on proteins. In particular, DNA and RNA libraries contain no positive charge, no functional groups (such as imidazole) that serve effectively as general acids and general bases near pH 7, and no non-aromatic hydrophobicity. Therefore, the aptamers and ribozymes obtained from standard IVS experiments with natural nucleic acids are limited in scope, especially when compared with the corresponding protein antibodies or enzymes.

AEGIS offers the opportunity to circumvent these limitations. The range of functionality available to the aptamer library pool can be expanded by introducing functionality onto nucleobases. The task targeted by the ONR was to develop an Expanded Combinatorial Selection (EXCSEL) tool that could generate receptors (for trinitrotoluene, for example) and catalysts (for degrading hazardous materials on Navy sites being closed and returned to civilian jurisdiction, for example), that could be applied to Navy needs.

Central to this goal is the development of DNA and RNA polymerases that will accept components of AEGIS. This has proven to be a challenging task. Our understanding of polymerases is inadequate to support an effort to directly design enzymes that accept these components. We have, of course, clones of various polymerases and the molecular biological tools needed to substitute specific amino acids in the sequence for others. The interactions between polymerases and components of AEGIS are subtle, however, and chemical theory is simply inadequate to predict what amino acid substitutions will generate the desired effects. Therefore, "directed evolution" (of the type used by Frances Arnold) experiments have been done with polymerases. These have been supported by the equipment purchased with this grant.

ACCOMPLISHMENTS:

Some 100 polymerases and naturally occurring variants of these have been examined, many using the equipment purchased under this grant. A proximity scintillation assay has been developed to support a "high throughput" screen of polymerases obtained via directed evolution. The first *in*

vitro selection experiments incorporating AEGIS were run using polymerases developed in this work. While the results of these experiments are still preliminary, it appears as if functionalization improves modestly the binding potential of an aptamer, and improves dramatically the catalytic potential of molecules obtained via in vitro selection.

CONCLUSIONS:
The equipment purchased under this grant has facilitated the ONR-funded project to develop *in vitro* selection methods for developing new catalysts, receptors, and ligands.

SIGNIFICANCE:
One of the "Holy Grails" in the molecular sciences is a method to generate receptors, ligands, and catalysts "on demand", without the need for design. Should these tools be successfully developed in this laboratory, the Navy will have a breakthrough technology for generating these. They can be then applied to develop analytical devices to detect explosives, biological warfare agents, and chemical warfare agents based on receptors, and the catalysts needed to degrade these.

PATENT INFORMATION:
none

AWARD INFORMATION:
Nolan Summer Award, University of Nebraska (March 1998)
Arun Gunthikonda Memorial Award Lecture, Columbia University (November, 1998)

PUBLICATIONS AND ABSTRACTS:
Jurczyk, S., Kodra, J. T., Rozzell, J. D., Jr., Benner, S. A. and Battersby, T.R., Synthesis of oligonucleotides containing 2'-deoxy*iso*guanosine and 2'-deoxy-5-methyl*iso*-cytidine using phosphoramidite chemistry. *Helv. Chim. Acta* **17**, 793-811 (1998)

Lutz, S., Benner, S. A. Simple one pot synthesis of a 2'-tritium labeled C-deoxynucleoside. *Bioorg. Med. Chem. Lett.* in press (1999)

Lutz, S., Burgstaller, P, Benner, S. A. An *In Vitro* Screening Technique for Polymerases that can Incorporate Modified Nucleotides. Pseudouridine as a Substrate for Thermostable Polymerases. *Nucl. Acids Res.* submitted.

Synthesis of Oligonucleotides Containing 2'-Deoxyisoguanosine and 2'-Deoxy-5-methylisocytidine Using Phosphoramidite Chemistry

by Simona C. Jurczyk[a]), Janos T. Kodra[b]), J. David Rozzell[a]), Steven A. Benner, and Thomas R. Battersby[a])*

Department of Chemistry, University of Florida, Gainesville, Florida 32611, USA

[a]) *Sulfonics Inc.*, 12085 Research Drive, Alachua, Florida 32615, USA
[b]) Department of Chemistry, Swiss Federal Institute of Technology, Universitätstrasse 16, CH-8092 Zürich

The synthesis of oligonucleotides containing 2'-deoxy-5-methylisocytidine and 2'-deoxyisoguanosine using phosphoramidite chemistry in solid-phase oligonucleotide synthesis is described. Supporting previous observations, the N,N-diisobutylformamidine moiety was found to be a far superior protecting group than N-benzoyl for 2'-deoxy-5-methylisocytidine. 2'-Deoxy-N^2-[(diisobutylamino)methylidene]-5'-(4,4'-dimethoxytrityl)-5-methylisocytidine 3'-(2-cyanoethyl diisopropylphosphoramidite) (**1c**) incorporated multiple consecutive residues during a standard automated synthesis protocol with a coupling efficiency > 99% according to dimethoxytrityl release. Extending coupling times of the standard protocol to ≥ 600 s using 2'-deoxy-N^6-[(diisobutylamino)methylidene]-5'-O-(dimethoxytrityl)-O^2-(diphenylcarbamoyl)isoguanosine, 3'-(2-cyanoethyl diisopropylphosphoramidite) (**7e**) led to successful incorporation of multiple consecutive 2'-deoxyisoguanosine bases with a coupling efficiency > 97% according to dimethoxytrityl release.

Introduction. – The *Watson-Crick* base pairing between two complementary oligonucleotide strands remains one of the most remarkable examples of molecular recognition. It follows two rules of complementarity: *i*) a large purine from one strand pairs with a small pyrimidine from the other; *ii*) H-bond donors (NH groups) of one base pair with H-bond acceptors (lone pairs of electrons on an O- or N-atom) of the other. Nature only partially exploits the potential of the *Watson-Crick* formalism, however. Structures for six base pairs can readily be written to conform to *Watson-Crick* geometry [1][2]. Therefore, it is possible, in principle, to have twelve independently replicating 'letters' in the nucleoside 'alphabet'.

One feasible non-standard base pair is formed between 2'-deoxyisocytidine (isoC$_d$, **1a**) and 2'-deoxyisoguanosine (isoG$_d$, **2a**), held together by three H-bonds like a G · C base pair. It was suggested three decades ago by *Rich* [3] that the isoG · isoC base pair might have been a component of primitive nucleic acids early in the development of life. In the late 1980's, the first experimental work was done to explore the oligonucleotide chemistry of this non-standard base pair, including its ability to be incorporated by enzymatic template-directed polymerization. Since then, several other laboratories have made major contributions to developing the chemistry and enzymology of the isoC · isoG base pair, including those of *Tor* and *Dervan* [4], *Horn et al.* [5], *Seela* and *Wei* [6] and *Switzer et al.* [7]. This work has shown, *inter alia*, that these bases can pair in both an antiparallel duplex as well as a parallel duplex, with similar stability to a guanine and cytosine base pair [5][8–10]. Further, a variety of RNA and DNA polymerases have been found that catalyze template-directed incorporation of this base pair into DNA [7].
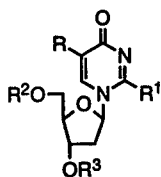
1a   R = R^2 = R^3 = H, R^1 = NH$_2$

b   R = H, R^1 = NHC(O)Ph, R^2 = (MeO)$_2$Tr,
R^3 = P(OCH$_2$CH$_2$CN)(N(i-Pr)$_2$)

c   R = Me, R^1 = NHC(O)Ph, R^2 = (MeO)$_2$Tr,
R^3 = P(OCH$_2$CH$_2$CN)(N(i-Pr)$_2$)

d   R = Me, R^2 = R^3 = H, R^1 = NH$_2$

e   R = Me, R^1 = NHC(O)Ph, R^2 = R^3 = H

f   R = Me, R^1 = NHC(O)Ph, R^2 = (MeO)$_2$Tr,
R^3 = H

g   R = Me, R^1 = N = CHNMe$_2$, R^2 = R^3 = H

h   R = Me, R^1 = N = CHN(i-Bu)$_2$, R^2 = R^3 = H

i   R = Me, R^1 = N = CHN(i-Bu)$_2$, R^2 = (MeO)$_2$Tr,
R^3 = H

j   R = Me, R^1 = N = CHN(i-Bu)$_2$, R^2 = (MeO)$_2$Tr,
R^3 = P(OCH$_2$CH$_2$CH)(N(i-Pr)$_2$)

l   R = Me, R^1 = NHC(O)Ph, R^2 = (MeO)$_2$Tr,
R^3 = P(O)H(OCH$_2$CH$_2$CN)

m   R = Me, R^1 = NH$_2$, R^2 = (MeO)$_2$Tr, R^3 = H
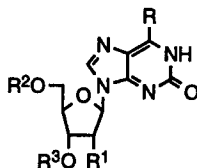
2a   R = NH$_2$, R^1 = R^2 = R^3 = H

b   R = N = CHN(i-Bu)$_2$, R^1 = R^2 = R^3 = H

c   R = N = CHN(i-Bu)$_2$, R^1 = H, R^2 = R^3 = Me$_3$Si

d   R = N = CHN(i-Bu)$_2$, R^1 = OH, R^2 = R^3 = H

e   R = N = CHN(i-Bu)$_2$, R^1 = OH,
R^2 = Si(i-Pr)$_2$OSi(i-Pr)$_2$OH, R^3 = H

f   R = N = CHN(i-Bu)$_2$, R^1 = OH,
R^2 = R^3 = (t-Bu)Me$_2$Si

g   R = NH$_2$, R^1 = OH, R^2 = R^3 = H

h   R = N = CHN(i-Bu)$_2$, R^1 = Me$_3$SiO,
R^2 = R^3 = Me$_2$Si

Despite the structural similarities that isoG and isoC share with the standard bases, it has not been trivial to convert procedures and protecting groups used for the automated solid-phase synthesis of standard oligonucleotides to be suitable for preparing oligonucleotides containing isoC and isoG. For example, *Switzer* and coworkers [9] reported that formamidine-protected 2′-deoxyisocytidine underwent rapid depyrimidination under conditions used routinely for the deprotection of oligonucleotides containing standard nucleobases. In contrast, *Horn et al.* [5] mentioned no such problem in their work. We found that isoC$_d$ as its benzoyl-protected derivative underwent deamination under basic oligonucleotide-deprotection conditions [7]. Likewise, *Strobel et al.* [11] reported that 5-methylisocytidine derivatives deaminated when subjected to these conditions. *Horn et al.* [5], however, reported no deamination upon basic deprotection of synthetic oligonucleotides using formamidine-protected 2′-deoxy-5-methylisocytidine phosphoramidites, although they did report depyrimidination 'precluding incorporation of more than a single disoMeC residue' into an oligonucleotide when the disoMeC (= me^{5}isoC$_d$) nucleobase was protected as a benzoate. In contrast, *Tor* and *Dervan* [4] reported neither depyrimidination nor deamination in oligonucleotide syntheses with the same benzoyl-protected me^{5}isoC$_d$.

Similar incongruitites can be found in the literature reporting the synthesis of oligonucleotides containing isoG. *Roberts et al.* [8][9] and *Sugiyama et al.* [10] mentioned no difficulties using standard solid-phase procedures to prepare oligonucleotides containing the phosphoramidite of 2′-deoxyisoguanosine protected as the N,N-formamidine and O^2-diphenylcarbamylate (dpc) ester. *Horn et al.* [5] reported, however, that dpc-protected isoG phosphoramidite 'may be too acid-labile for practical use'. *Seela* and *Wei* [6] noted that solid-phase synthesis with 2′-deoxyisoguanosine using phosphoramidites and

phosphonates gave low yields with consecutive incorporations, and suggested that a modified phosphonate synthesis be used for oligomers containing multiple consecutive isoG residues.

In view of the importance of the non-standard base pair in a variety of areas, most interestingly in DNA-based diagnostics [5][12], it is timely to report a detailed study containing full experimental recipes describing the chemistry for the synthesis of oligonucleotides containing $isoC_d$ and $isoG_d$. In this work, we develop syntheses and evaluate the utility of various protecting groups for phosphoramidites of $isoC_d$ and $isoG_d$ in automated synthesis, exploring and resolving some of the incongruities noted above.

Results and Discussion. – *Isocytidine.* Because of reports that deamination was suppressed by a 5-methyl substituent in isoC [4] (*Scheme 1*), this work focused on protected variants of this derivative. Thus, N^2-benzoyl-2'-deoxy-5'-O-(4,4'-dimethoxytrityl)-5-methylisocytidine 3'-(2-cyanoethyl diisopropylphosphoramidite) (**1c**) was prepared following the general procedure of *Tor* and *Dervan*, with two improvements that significantly increased the yields. First, use of a *Parr* pressure reactor for the treatment of 2,5'-anhydrothymidine (**3**) with ammonia resulted in an improved, nearly quantitative yield of 2'-deoxy-5-methylisocytidine (**1d**). The synthetic results were also improved by chromatographic purification (silica gel) of the tritylated derivative **1f** immediately following its formation from **1e**, without storage. After modification, the overall yield for the three steps producing **1c** from 2'-deoxy-5-methylisocytidine (**1d**) was 36%.

When used in a standard oligonucleotide synthesis, N^2-benzoyl-2'-deoxy-5'-O-(4,4'-dimethoxytrityl)-5-methylisocytidine 3'-(2-cyanoethyl diisopropylphosphoramidite) (**1b**) did not perform satisfactorily in our hands. The trityl release following coupling of the non-standard nucleotide derivative dropped precipitously, and very little full-length product was detected upon completion of the synthesis. An analogous observation was made by *Horn et al.* [5]. These authors proposed that the low coupling yield arose due to depyrimidination, and they suggested the use of formamidine protection to avoid this problem.

Depyrimidination occurs during the acidic detritylation steps employed in automated synthesis, so the efficacies of several detritylation reagents on the model compound 5'-O-(4,4'-dimethoxytrityl)thymidine ([(MeO)$_2$Tr]T$_d$) were investigated. Three detritylation reagents proposed for oligonucleotide synthesis were examined: CCl_3COOH, $ZnBr_2$ [13], and ceric ammonium nitrate [14]. CCl_3COOH (1%) in 1,2-dichloroethane and $ZnBr_2$ in MeNO$_2$/PhNO$_2$ 4:1 both detritylated [(MeO)$_2$Tr]T$_d$ almost instantaneously at room temperature. Interestingly, a substantial amount of (MeO)$_2$Tr-protected thymidine remained after 15 min when ceric ammonium nitrate in wet MeCN or wet MeCN/DMF was used.

N^2-Benzoyl-2'-deoxy-5-methylisocytidine (**1e**) and the formamidine-derived 2'-deoxy-N^2-[(dimethylamino)methylidene]-5-methylisocytidine (**1g**) were treated with the three detritylating reagents at room temperature (see *Table*). The resulting mixtures were quenched and then analyzed on reversed-phase HPLC, monitoring at 260 nm. The N^2-benzoyl derivative **1e** degraded completely in less than 15 min in the CCl_3COOH and $ZnBr_2$ solutions. In solutions containing ceric ammonium nitrate, **1e** decomposed more slowly, but the detritylation of the corresponding trityl derivative **1f** was also slower, with depyrimidination faster than detritylation. A larger sample of the only decomposition

Scheme 1
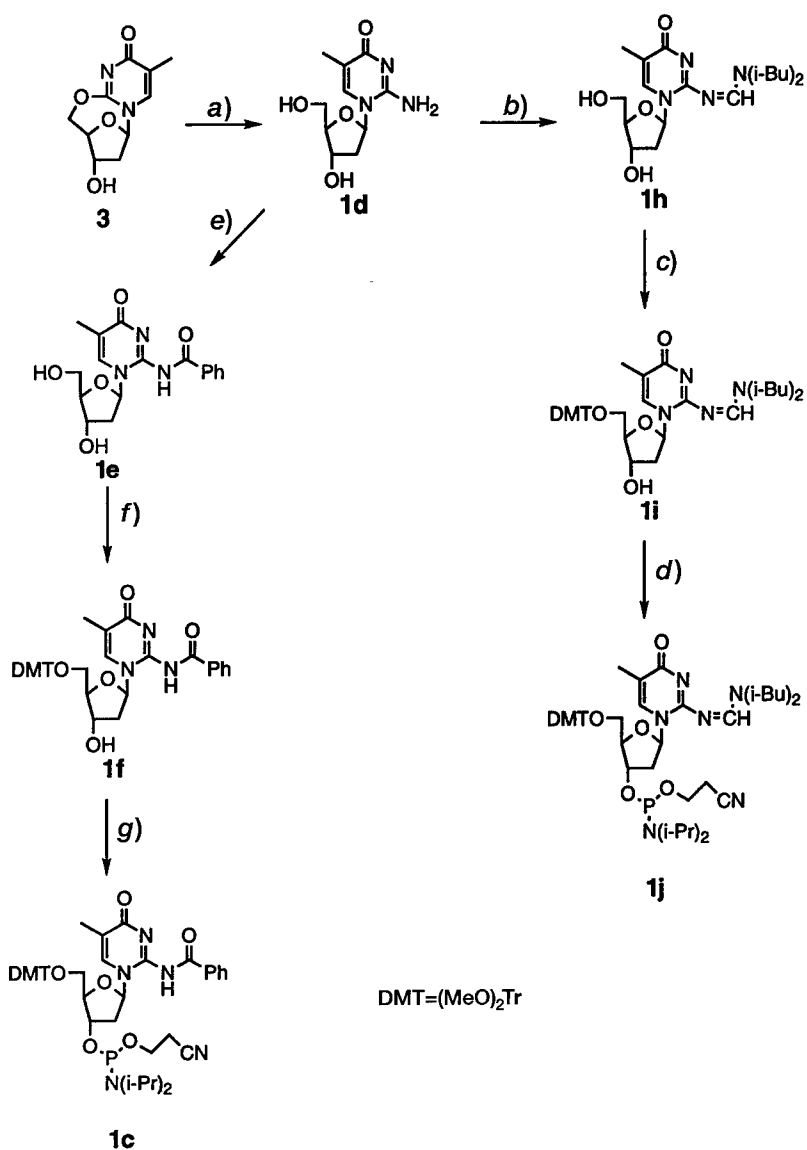
a) NH$_3$, MeOH; 98%. *b*) (i-Bu)$_2$NCH(OMe)$_2$, DMF; 92%. *c*) (MeO)$_2$TrCl, pyridine, 4-(dimethylamino)pyridine, DMAP; 52%. *d*) 2-Cyanoethyl diisopropylphosphoroamidochloridite, (i-Pr)$_2$NEt, CH$_2$Cl$_2$; 79%. *e*) PhCOCl, pyridine, DMAP; 66%. *f*) (MeO)$_2$TrCl, pyridine, DMAP; 73%. *g*) 2-Cyanoethyl diisopropylphosphoramido-chloridite, (i-Pr)$_2$NEt, CH$_2$Cl$_2$; 74%.

product absorbing at 260 nm observed after exposure of **1e** to the detritylation conditions was generated with CCl$_3$COOH and confirmed as 5-methylisocytosine by ^{1}H-NMR.

Table. *Approximate Time Needed for Complete Depyrimidination of 2'-Deoxy-5-methylisocytidine Derivates* (detection level *ca.* 1%)

	1% CCl_3COOH/ 1,2-dichloroethane (r.t.)	$ZnBr_2$ in nitrobenzene/ nitromethane 1:4 (r.t.)	$Ce^{IV}(NH_4)_2(NO_3)_6$ (1.2 mM) in $MeCN/DMF/H_2O$ 6:2:0.05 (reflux)
1e	<15 min	<15 min	150 min
1g	36 h	7 days	–

The corresponding formamidine-protected N^2-[(dimethylamino)methylidene] derivative **1g** was considerably more stable. After 1 h in the CCl_3COOH solution, only *ca.* 1% of the compound was degraded (UV). In the presence of $ZnBr_2$, no depyrimidination of **1g** was detected after 1 h at room temperature. Even after a further 16 h at 40°, only *ca.* 2% degradation was detected.

To evaluate the stability of **1g** with respect to deamination, the compound was exposed to standard cleavage/deprotection conditions used in automated synthesis. A small amount of **1g** was placed in 25% aqueous ammonium hydroxide and heated at 60° for 16 h. The solvent was removed under vacuum, the residue dissolved in water, and the mixture resolved by HPLC. Prior to treatment, UV analysis of an HPLC chromatogram showed **1g** to account for 99% of the material absorbing at 260 nm. Following treatment, 2'-deoxy-5-methylisocytidine (**1d**) accounted for 96% of the material. Thymidine (perhaps 2%, verified by addition of authentic thymidine) was the only detectable impurity, corresponding to the deamination product of **1d**.

These results show that N^2-benzoyl-2'-deoxy-5-methylisocytidine (**1e**) is too acid-labile for practical use; it was even labile under non-protic detritylation conditions. The successful use of this compound in oligonucleotide synthesis by *Tor* and *Dervan* [4] might be ascribed to the fact that the oligonucleotide prepared contained a single isoC residue close to the 5'-end of the oligomer, and therefore did not need to survive many cycles of treatment under detritylation conditions.

In contrast, the formamidine-protected 5-methylisocytidine **1g** appears suitable for phosphoramidite automated synthesis, displaying only small amounts of depyrimidination and (perhaps) deamination. The dimethylformamidine derivative **1g** used in the stability study, however, has been reported as labile under DNA-synthesis conditions [15]. Therefore, 2'-deoxy-N^2-[(diisobutylamino)methylidene]-5'-O-(4,4'-dimethoxytrityl)-5-methylisocytidine 3'-(2-cyanoethyl diisopropylphosphoramidite) (**1j**), bearing the more stable diisobutylformamidine protecting moiety, was synthesized in three steps from 2'-deoxy-5-methylisocytidine (**1d**). Reaction with N,N-diisobutylformamidine dimethyl acetal gave the easily-purified N^2-protected isocytidine **1h** after 1 h (*Scheme 4*). The fully protected target phosphoramidite **1j** was readily isolated in an overall yield of 38% for the three steps *via* **1h** and **1i** starting with **1d**.

Additionally, a method adapted from a procedure for recycling unused phosphoramidites from an automated synthesizer [16], proceeding through intermediates **1l** and **1m**, was developed to convert existing benzoyl-protected phosphoramidite into formamidine-protected phosphoramidite (**1c** → **1l** → **1f** → **1m**; see *Exper. Part*).

Phosphoramidite **1j** was examined for utility for synthesizing oligonucleotides containing isoC. Although $ZnBr_2$ was (from the studies above with **1g**) a superior detritylat-

ing reagent without depyrimidination, it was decided to first use standard detritylating conditions (CCl_3COOH), which also performed well in the stability study. The three oligonucleotides **4–6** were synthesized to test the efficacy of phosphoramidite **1j**.

A PAGE autoradiogram and an anion-exchange HPLC (*Fig. 1*) of the crude synthetic oligonucleotide **4** showed small amounts of failure products (albeit greater than seen with standard bases). Nevertheless, the synthesis yielded a satisfactory amount of full-length oligonucleotide **4**. The synthesis of crude oligonucleotide **5** gave similar results. After deprotection and desalting, the crude mixtures containing **4** or **5** were purified by ion-exchange HPLC, followed by reversed-phase HPLC. To verify the identity of the major product, the purified oligonucleotides **4** and **5** were digested by snake-venom phosphodiesterase and alkaline phosphatase and analyzed by reversed-phase HPLC. The component nucleosides found were consistent with the expected results for digestion of **4** and **5**, respectively (*Fig. 2*). Notably, no significant peak arising from T_d was present in the chromatograms, indicating that deamination did not occur in the oligonucleotides to the same extent as observed in syntheses with nucleoside **1b**, or even to the extent observed above with the formamidine-protected monomer nucleoside **1g**.

5'-d(A-iC-iC-iC-iC-iC-iC-iC-iC-iC-C) (**4**)

5'-d(A-A-A-A-A-iC-iC-iC-iC-iC-A-A-A-A-A) (**5**)

5'-d(T-T-T-T-T-iC-iC-iC-iC-iC-T-T-T-T-T) (**6**)

5'-d(T-T-T-T-T-iG-iG-iG-iG-iG-T-T-T-T-T) (**10**)

5'- d(C-A-C-A-A-iG-A-iG-A-iG-A-A-C-A) (**11**)

Oligonucleotide syntheses are frequently followed by quantifying the amount of trityl cation released. The $isoC_d$-containing oligonucleotide **6** was synthesized to evaluate coupling during synthesis by analyzing the integrated absorbance of released $(MeO)_2Tr$ protecting group at each added base. The efficiency of synthesis of the first part of oligonucleotide **6** was determined from couplings 2–4, the second part (containing $isoC_d$) from couplings 5–9, and the final part from couplings 10–14. The coupling efficiency of the $isoC_d$-derived phosphoramidites was 99.5%, comparable to the 99.9% efficiency for coupling in the other two sections containing only T_d.

Isoguanosine. Synthesis of oligonucleotides containing 2'-deoxyisoguanosine has been previously reported by *Switzer et al.* [7] using the 2'-deoxy-N^6-[(dibutylamino)-methylidene]-5'-O-(4,4'-dimethoxytrityl)-O^2-([2-4-nitrophenyl)ethyl]isoguanosine 3'-(2-cyanoethyl diisopropylphosphoramidite) monomer. To avoid the use of 1,8-diazabicyclo[5.4.0]undec-7-ena (DBU) in the deprotection procedure, the diphenylcarbamoyl moiety has been suggested for protecting the 2-O-position [5][9]. This protecting group has been used successfully for incorporation of 2'-deoxyisoguanosine using phosphonate chemistry [6], and evidently helps stabilize the N-glycoside bond from acid-catalyzed hydrolysis [6].

Thus, 2'-deoxy-N^6-[(diisobutylamino)methylidene]-5'-O-(4,4'-dimethoxytrityl)-O^2-(diphenylcarbamoyl)isoguanosine 3'-(2-cyanoethyl diisopropylphosphoramidite) (**7e**) was first synthesized in 7 steps (*Scheme 2*) from 2'-deoxy-2-aminoadenosine (**7a**). Although this synthesis is satisfactory for small amounts of the target phosphoramidite **7e**,
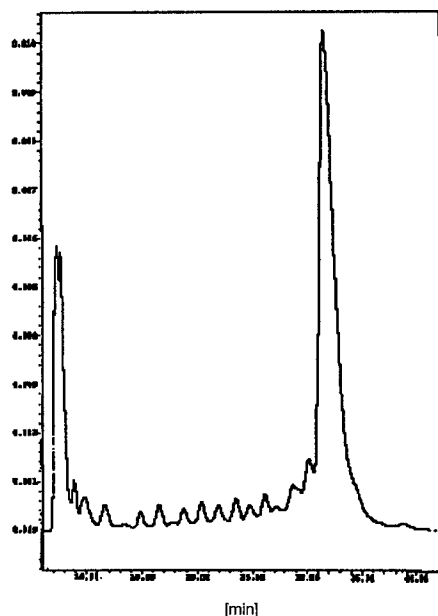
[min]

Fig. 1. *Anion-exchange HPLC of the crude synthetic oligonucleotide* **4**
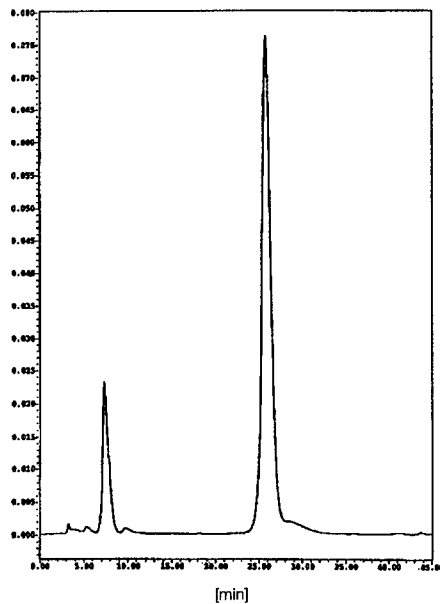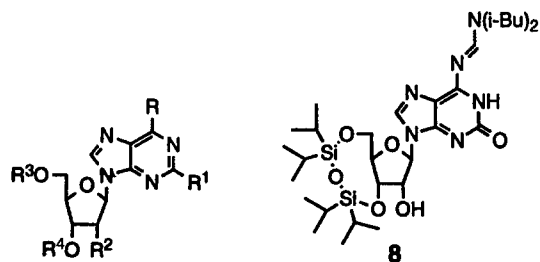

[min]

Fig. 2. *Reversed-phase HPLC of the enzymatic digestion products obtained from* **5**. The product of deamination of isoC$_d$, T$_d$, is present in less than 1%. isoC$_d$ (8.0 min), A$_d$ (26 min).


7a R = R^1 = NH$_2$,
  R^2 = R^3 = R^4 = H
 b R = N = CHN(i-Bu)$_2$,
  R^1 = OC(O)NPh$_2$,
  R^2 = H, R^3 = R^4 = Me$_3$Si
 c R = N = CHN(i-Bu)$_2$,
  R^1 = OC(O)NPh$_2$, R^2 = R^3 = R^4 = H
 d R = N = CHN(i-Bu)$_2$, R^1 = OC(O)NPh$_2$,
  R^2 = H, R^3 = (MeO)$_2$Tr, R^4 = H
 e R = N = CHN(i-Bu)$_2$, R^1 = OC(O)NPh$_2$,
  R^2 = H, R^3 = (MeO)$_2$Tr, R^4 = P(OCH$_2$CH$_2$CN)(N(i-Pr)$_2$)
 f R = R^1 = NH$_2$, R^2 = OH, R^3 = R^4 = H
 g R = N = CHN(i-Bu)$_2$, R^1 = OC(O)NPh$_2$,
  R^2 = Me$_3$SiO, R^3 = R^4 = Me$_3$Si
 h R = N = CHN(i-Bu)$_2$, R^1 = OC(O)NPh$_2$,
  R^2 = OH, R^3 = R^4 = H

9a R= OH
 b R= OC(S)OPh
 c R=H
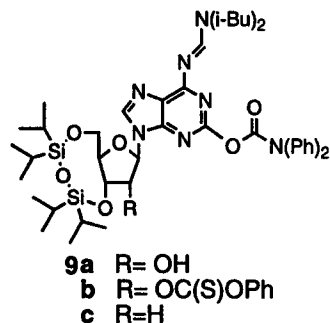
we required a procedure amenable to large-scale work. Unfortunately, the starting material **7a** in this route is quite expensive. Consequently, an alternative route starting with the relatively inexpensive ribose compound, 2-aminoadenosine (**7f**), was designed.
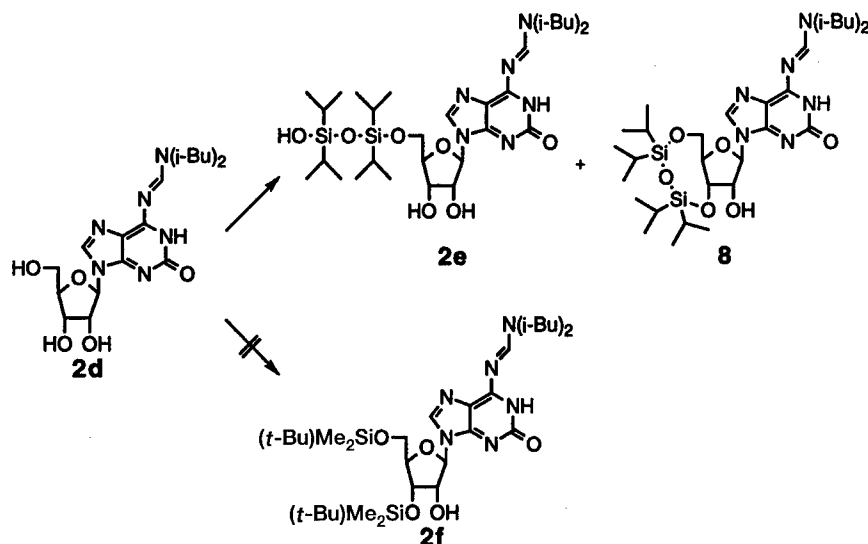
Scheme 2
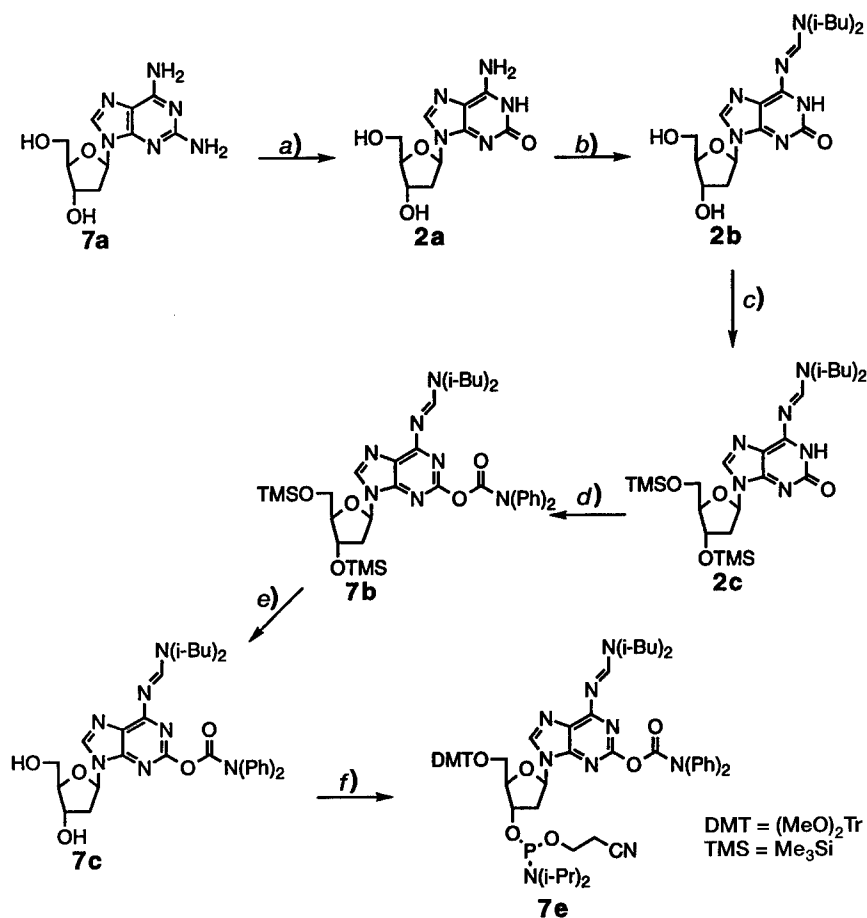
a) AcOH, NaNO$_2$, H$_2$O; 81%. b) (i-Bu)$_2$NCH(OMe)$_2$, DMF; 96%. c) Me$_3$SiCl, Et$_3$N, THF; 77%. d) Ph$_2$NCOCl, pyridine, (i-Pr)$_2$NEt; 75%. e) THF; 89%. f) 1. (MeO)$_2$TrCl, pyridine, DMAP; 76%; 2. 2-cyano-ethyl diisopropylphosphoramidochloridite, (i-Pr)$_2$NEt, CH$_2$Cl$_2$; 74%.

Originally the intermediate **8** was planned in the synthesis of **7e** from **7f**. Reaction of **2d** with 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane gave a *ca.* 1:1 mixture (^{1}H-NMR) of **8** and **2e** (*Scheme 3*), however. Problems forming the eight-membered cycle bridging the O-atoms at C(3′) and C(5′) have been observed in other nucleosides. Silylations at the 3′-O- and 5′-O-positions to give desired product, as well as at an unprotected carbonyl group of the base and the 5′-O-position to give side product, have been observed for unprotected isoG$_d$ [17] and for *N*-protected isoC [18]. Compounds **8** and **2e** were separable, but the resulting poor yield was impracticable. Next, selective 3′,5′-di-O-silylation of

2d to **2f** was tried (*Scheme 3*), using a literature procedure for similar reactions [19]; however, NMR analysis of the product revealed a complex mixture of several inseparable silylated derivatives.
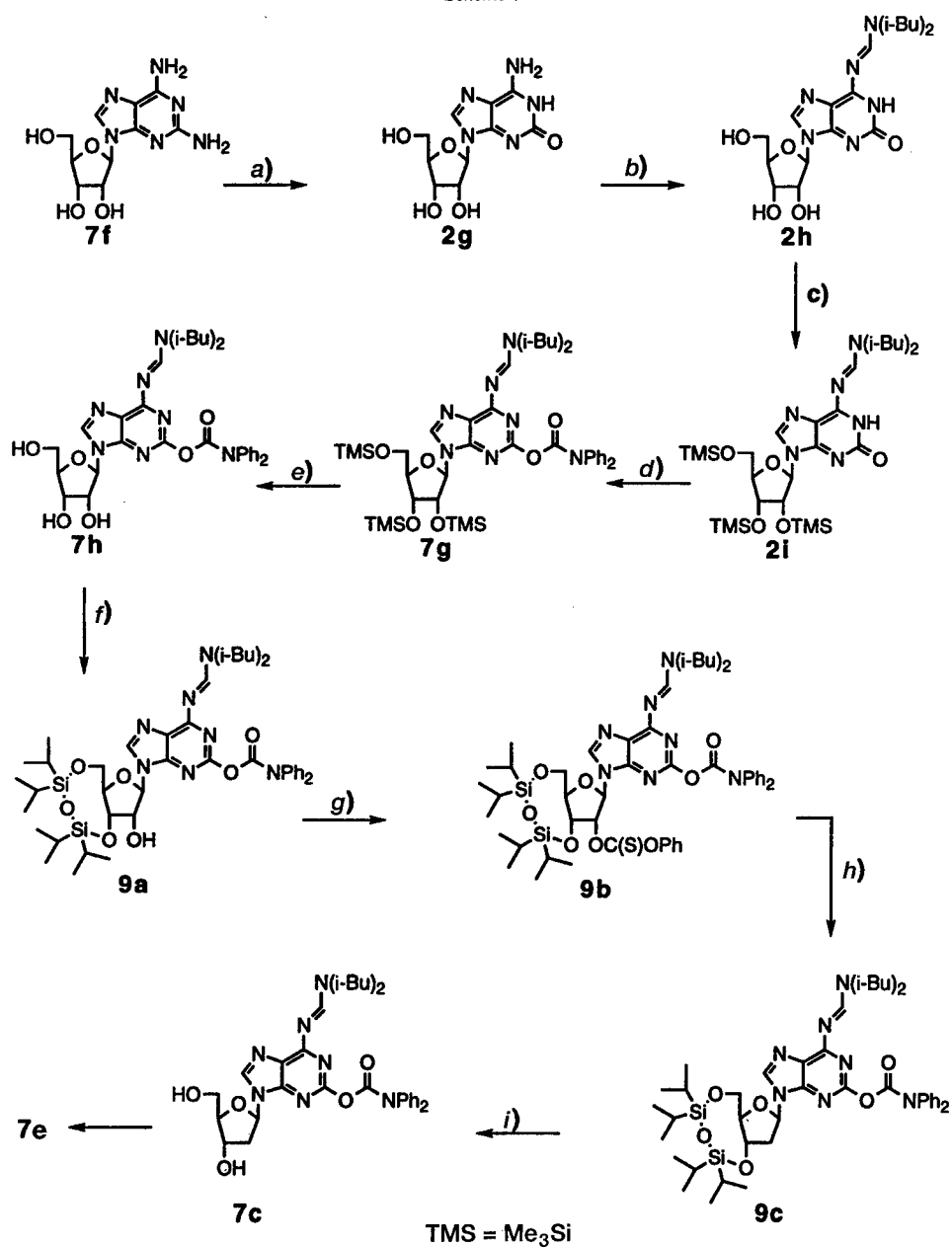
Scheme 3

A slightly longer route was then developed to synthesize **7e** from the ribose derivative **7f** (*Scheme 4*). The intermediate **2g** was purified with much greater ease than the corresponding 2'-deoxy compound **2a** (see *Scheme 3*), and not all synthetic intermediates required isolation. An overall yield of 13% for 11 steps, producing **7e** from the inexpensive starting material **7f** *via* the silylated derivatives **9a–c** without side reactions or laborious purifications, make this route suitable for large-scale preparation.

Coupling efficiency with phosphoramidite **7e** was evaluated in the same manner as described for the isoC$_d$-derived phosphoramidite **1j** above. Incorporation of **7e** into oligonucleotide **10** yielded little or no target oligomer using standard coupling times. The situation was greatly improved with extended coupling times (*Fig. 3*). It proved necessary to extend the coupling times to at least 600 s for each cycle incorporating an isoG$_d$ residue to obtain reasonable coupling efficiency. The increase in full-length target oligomer with extended coupling times was verified by ion-exchange HPLC of the crude products. Despite high coupling yields by trityl release, the HPLC indicate significantly more failure products than seen with standard bases, although the full-length oligonucleotide **10** was clearly the major product of the synthesis. The longer coupling time (600 s) needed for the incorporation of the isoG$_d$ residue compared to conventional DNA synthesis (*ca.* 90 s) lacks an explanation. In any case, the coupling time required for oligonucleotide synthesis containing 2'-deoxyisoguanosine is comparable to the coupling time employed in oligoribonucleotide synthesis [20].

Oligonucleotide **11** was synthesized with extended coupling times (900 s) using phosphoramidite **7e**. An anion-exchange HPLC of the crude product indicated that although the trityl-release-monitored coupling efficiency was > 95%, quite significant amounts of

Scheme 4

TMS = Me₃Si

a) AcOH, NaNO₂, H₂O; 91%. *b*) (i-Bu)₂NCH(OMe)₂, DMF; 90%. *c*) Me₃SiCl, Et₃N, THF; 96%.
d) Ph₂NCOCl, pyridine, (i-Pr)₂NEt; 77%. *e*) Bu₄NF, THF; 91%. *f*) 1,3-Dichloro-1,1,1,3,3,3-tetraisopropyldisilo-
xane, pyridine; 92%. *g*) PhOC(S)Cl, pyridine, DMAP, CH₂Cl₂; 94%. *h*) Bu₃SnH, AIBN, toluene; 73%.
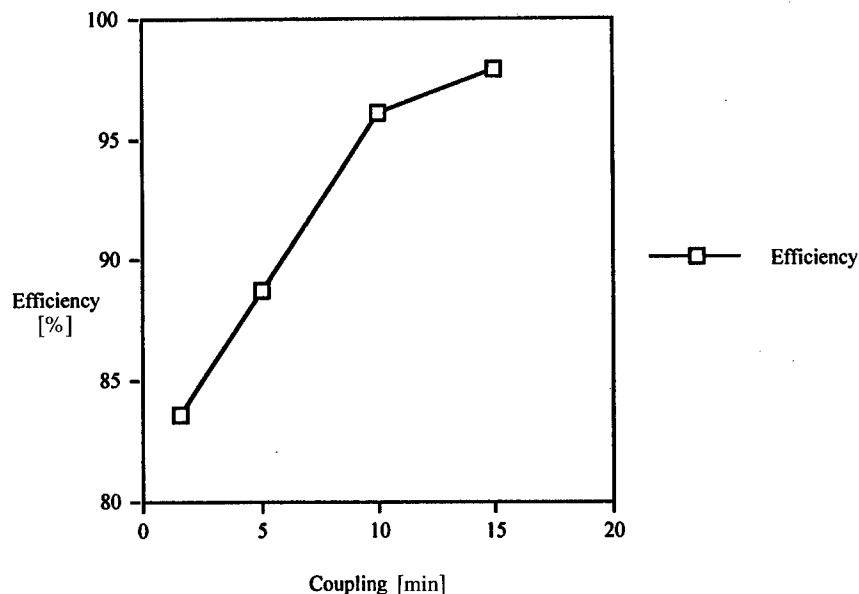*i*) Bu₄NF, THF.

isoG$_d$ Couplings

Fig. 3. *The coupling efficiency* [%] *by trityl-release monitoring of isoG$_d$ phosphoramidite* **7e** *vs. coupling time in syntheses of* **10**

failure products were generated. The crude product was purified by anion-exchange HPLC, followed by reversed-phase HPLC, and the purified oligonucleotide **11** was enzymatically digested as described above. The reversed-phase HPLC of the digestion products are consistent with the expected component nucleosides of **11** (*Fig. 4*).

Experimental Part

General. The 2-aminoadenosine (**7f**) and 2-amino-2'-deoxyadenosine (**7a**) were purchased from *RI Chemicals, Inc.* (Orange, CA), 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane from *Farchan Laboratories, Inc.* (Gainesville, FL), 1*H*-imidazole from *Kodak,* and all other chemicals from *Aldrich Chemical* or *Fisher Scientific;* they were used without further purification. All materials to make buffer solutions, including K_2HPO_4, KCl, Et_3N, *Tris* (= tris(hydroxymethyl)aminomethane (= 2-amino-2-(hydroxymethyl)propane-1,3-diol)), HCl, and AcOH were from *Fisher,* except for KH_2PO_4, which was from *Mallinckrodt.* TsCl was freshly recrystallized from petroleum ether. Solvents were dried over 4-Å molecular sieves. CH_2Cl_2 used in the synthesis and workup of phosphoramidites was freshly distilled from K_2CO_3. All reactions were carried out under dry Ar in an oven-dry glass system. 'Evaporation' refers to removal of volatile solvents with a membrane pump. Column chromatography (CC): silica gel (230–425 mesh). TLC: silica gel TLC plates from *Whatman,* visualization by staining with a Ce/Mo reagent (2.5% phosphormolybdic acid, 1% $Ce^{IV}(SO_4)_2 \cdot 4H_2O$, 6% H_2SO_4 in H_2O) and heating. HPLC: solvents from *Fisher Scientific* (HPLC grade) and *Milli*-Q-purified water (*Millipore Corp.*); solvents filtered through a 0.45 µm *Whatman* nylon filter just before use; *Waters Alliance* system with a *486* tunable absorbance detector or *Waters* system consisting of a *600S* controller, *616* pump, and a *996* photodiode array detector; for anion exchange,
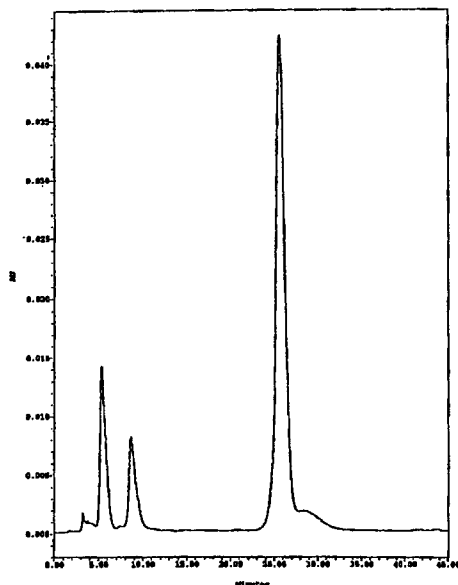
Fig. 4. *Reversed-phase HPLC of the enzymatic digestion products of* **11**. C_d (5.5 min), $isoG_d$ (9.0 min), A_d (26 min).

Alltech Associates column (*Macrosphere 300A WAX*, 7 μm, 4.6 × 250 mm); for reversed-phase chromatography, two *Waters* columns (*Nova-Pak HR* C_{18} *60A*, 6 μm, 7.8 × 300 mm, or *Nova-Pak* C_{18} *60A*, 4 μm, 3.9 × 150 mm). NMR Spectra: *Varian-XL-300* spectrometer at 300 MHz referenced to $SiMe_4$ (^1H), at 75.4 MHz and referenced to $CDCl_3$ (^{13}C) and at 121.4 MHz with H_3PO_4 as standard (^{31}P).

2,5′-Anhydrothymidine (**3**). The reported procedure [21] with a reaction time of 24 h resulted in formation of decomposition products; the reaction was nearly complete after 3.5 h (TLC (CH_2Cl_2/MeOH 3:1): R_f 0.42). Further, much solid product was discarded in the filtration of the hot reaction mixture following the literature workup. The procedure was improved as follows: According to [21], 5′-*O*-(*p*-tolylsulfonyl)thymidine (75.7 mmol, 30.0 g) was prepared and then suspended in dry MeCN (450 ml). DBU (1.1 equiv., 83.32 mmol, 12.66 g, 12.44 ml) was added and the mixture refluxed for 3.5 h. Then more DBU (0.1 equiv, 1.2 ml) was added, the mixture refluxed for 1 more h, then cooled to r.t., and filtered, and the filtrate collected. The solid was refluxed with MeOH (150 ml) for 5 min and the mixture filtered while hot. The solid was again refluxed with MeOH (50 ml) for 5 min and filtered. The combined filtrates were evaporated and triturated with MeCN. The solid product (9.86 g) was dried *in vacuo* overnight. The MeCN filtrate obtained above was evaporated, acetone (150 ml) added, and the mixture kept for several hours. The product (1.60 g) was isolated as long needles and dried *in vacuo*. The products were combined (11.5 g, 68%). ^1H-NMR (D_2O/CD_3OD): 1.74 (*s*, Me–C(5)); 2.21–2.52 (*m*, 2H–C(2′)); 4.00 (*d*, 1H–C(5′)); 4.28 (*m*, H–C(4′)); 4.38 (*d*, 1H–C(5′)); 4.50 (*m*, H–C(3′)); 5.84 (*d*, H–C(1′)); 7.75 (*s*, H–C(6)). ^{13}C-NMR (D_2O/CD_3OD): 13.4; 43.2; 73.0; 76.5; 87.5; 95.7; 119.7; 142.1; 159.1; 175.5.

2′-Deoxy-5-methylisocytidine (**1d**). The yield of the reported procedure [4] was greatly improved by the use of a 600-ml pressure reactor (heater, mechanical stirrer; *Parr* instrument) which also allowed the reaction to be easily scaled up: Compound **3** (43.0 mmol, 9.63 g) was added to MeOH saturated at 0° with NH_3 (250 ml), and the mixture was placed in the pressure reactor and heated to 103° for 4 h. The reactor was cooled to 0° and opened, the solvent evaporated and the solid product taken up in MeCN and filtered. The white powder (10.3 g, 99%) was dried *in vacuo* overnight. ^1H-NMR (D_2O/CD_3OD): 1.61 (*s*, Me); 2.20 (*m*, 2H–C(2′)); 3.58 (*m*, 2H–C(5′)); 3.80 (*m*, H–C(4′)); 4.21 (*m*, 1H–C(3′)); 5.67 (*t*, H–C(1′)); 7.39 (*s*, H–C(6)). ^{13}C-NMR (D_2O/CD_3OD): 13.7; 39.2; 61.7; 71.0; 87.9; 89.7; 115.9; 137.3; 155.9; 174.9.

N-Benzoyl-2′-deoxy-5-methylisocytidine (**1e**). A modified literature procedure [4] was inimicable to large-scale synthesis because of the large excess of PhCOCl required and the associated cumbersome purification; the excess of benzoyl chloride required could be reduced by addition of 4-(dimethylamino)pyridine (DMAP): Compound **1d** (1.77 mmol, 4.27 mg) was dissolved in anh. pyridine (15 ml) and DMAP (0.5 equiv., 0.886 mmol, 108 mg).

PhCOCl (5 equiv., 8.86 mmol, 1.03 ml) and Et$_3$N (5 equiv., 8.86 mmol, 1.23 ml) were added at 0°. The mixture was stirred at r.t. for 45 min (TLC (CHCl$_3$/MeOH 9:1): R_f 0.87). Aq. NaHCO$_3$ soln. was added at 0°, adjusting the pH to 7. The soln. was stirred at r.t. for 30 min and then extracted with CHCl$_3$ (3 × 50 ml). The combined org. layers were washed (NaHCO$_3$ soln. and H$_2$O), dried (Na$_2$SO$_4$), and evaporated. The residue was dissolved in pyridine/MeOH 3:1 (9 ml), and a soln. of 9 equiv. NaOH (16.0 mmol, 638 mg) in H$_2$O/MeOH 1:1 (20 ml) was added at 0°. After 10 min stirring at r.t., phosphate buffer soln. (5 ml) was added and the mixture evaporated. The residue was purified by CC (silica gel, CHCl$_3$/MeOH 9:1): 403 mg (66%) of **1e**. White solid. ^{1}H-NMR (D$_2$O/CD$_3$OD): 1.75 (s, Me); 2.08, 2.40 (2m, 2H–C(2′)); 3.63 (m, 2H–C(5′)); 3.92 (m, H–C(4′)); 4.28 (m, H–C(3′)); 6.58 (m, H–C(1′)); 7.20–7.40 (m, 3 H, Ph); 7.82 (s, H–C(6)); 7.92 (d, 2 H, Ph).

N^2-Benzoyl-2′-deoxy-5′-(4,4′-dimethoxytrityl)-5-methylisocytidine (**1f**). The reported procedure [4] (12 h at r.t.) gave mainly unreacted starting material **1e**, even after a 20 h reaction time; better results were achieved at higher temp., with increased excess of (MeO)$_2$TrCl, and by extraction with AcOEt instead of Et$_2$O. Thus, to a soln. of **1e** (1.54 mmol, 530 mg) in dry pyridine (20 ml) and DMAP (0.25 equiv., 0.384 mmol, 46.8 mg), Et$_3$N (2 equiv., 3.1 mmol, 310 mg, 0.43 ml) and (MeO)$_2$TrCl (1.8 equiv., 2.77 mmol, 936 mg) were added. The mixture was stirred at 40° for 3.5 h and at r.t. for 1 h. MeOH (3 ml) was added to quench the reaction, followed by aq. NaHCO$_3$ soln. (50 ml). The mixture was extracted with AcOEt (1 × 50 ml, 3 × 30 ml), the extract dried (Na$_2$SO$_4$) and evaporated, and the residue purified by CC (CH$_2$Cl$_2$/MeOH 98.5:1.5; R_f 0.27): yellowish foam (726 mg, 73%). ^{1}H-NMR (CDCl$_3$): 1.49 (s, Me–C(5)); 2.42, 2.70 (2m, 2H–C(2′)); 3.50 (dd, 2H–C(5′)); 3.74 (s, 2 MeO); 4.22 (m, H–C(4′)); 4.68 (m, H–C(3′)); 6.81 (dd, 4 H, (MeO)$_2$Tr); 6.91 (t, H–C(1′)); 7.20–7.46 (m, 12 H, Ph, (MeO)$_2$Tr); 7.88 (s, H–C(6)); 8.18 (d, 2 H, Ph). ^{13}C-NMR (CDCl$_3$): 12.3; 41.7; 55.1; 55.2; 63.0; 71.3; 71.4; 86.2; 86.3; 86.9; 113.2; 114.7; 127.1; 127.91; 127.97; 128.0; 129.4; 130.0; 132.1; 135.1; 135.2; 135.5; 144.2; 152.5; 158.6; 161.0; 177.4.

N^2-Benzoyl-2′-deoxy-5′-O-(4,4′-dimethoxytrityl)-5-methylisocytidine 3′-O-(2-Cyanoethyl Diisopropylphosphoramidite) (**1c**). The reaction time of the reported procedure [4] was shortened to avoid formation of decomposition products, and the product was purified by reprecipitation to remove traces of phosphoramidite by-products: A soln. of **1f** (co-evaporated from MeCN; 6.16 mmol, 3.98 g) in CH$_2$Cl$_2$ was cooled to 0°, and (i-Pr)$_2$EtN (5.5 equiv., 33.9 mmol, 4.37 g, 5.89 ml) and 2-cyanoethyl diisopropylphosphoramidochloridite (1.4 equiv., 8.62 mmol, 2.04 g, 2.01 ml) were added via syringe. The mixture was stirred for 10 min at 0° and then 30 min at r.t. Aq. NaHCO$_3$ soln. (100 ml) was added and the mixture extracted with CH$_2$Cl$_2$/Et$_3$N 98:2 (3 × 70 ml). The org. layer was dried (Na$_2$SO$_4$) and evaporated and the residue purified by CC (CH$_2$Cl$_2$/AcOEt/Et$_3$N 7:3:1; R_f 0.84 and 0.86, 2 diastereoisomers). The product was dissolved in CH$_2$Cl$_2$ (10 ml), this soln. added slowly dropwise to strongly stirred hexane (700 ml) at r.t., the mixture decanted, and the solid product dried in vacuo: pale-yellow foam (3.86 g, 74%). ^{1}H-NMR (CDCl$_3$): 1.08–1.24 (m, 2 Me$_2$CH); 1.47, 1.51 (2s, Me–C(5)); 2.34–2.81 (m, 2H–C(2′), CH$_2$CN); 3.37 (dd), 3.54–3.89 (m, 2H–C(5′), 2 Me$_2$CH, OCH$_2$CH$_2$CN); 3.78 (d, 2 MeO); 4.27 (m, H–C(4′)); 4.69 (m, H–C(3′)); 6.84 (dd, 4 H, (MeO)$_2$Tr); 6.91 (m, H–C(1′)); 7.22–7.54 (m, 12 H, (MeO)$_2$Tr, Ph); 7.91, 7.95 (2s, H–C(6)); 8.20 (m, 2 H, Ph). ^{13}C-NMR (CDCl$_3$): 12.3; 20.2; 24.5; 40.6; 43.1; 43.4; 55.1; 58.0; 62.7; 73.3; 85.7; 86.0; 86.9; 113.2; 114.9; 117.3; 117.4; 127.1; 127.9; 128.0; 129.3; 130.0; 132.0; 135.1; 135.2; 137.0; 137.1; 144.1; 152.7; 152.8; 158.7; 160.7; 177.2. ^{31}P-NMR (CDCl$_3$): 149.24; 150.28.

N,N-Diisobutylformamide Dimethyl Acetal. According to a slightly modified literature procedure [22][23], a mixture of *N,N*-dimethylformamide dimethyl acetal (754 mmol, 100 ml) and diisobutylamine (754 mmol, 131 ml) was refluxed in an oil bath at 105–115° for 5 days. After evaporation of the volatile compounds, the product was distilled under reduced pressure using a 35-cm column: 53.3 g (35%). B.p. 92°/17 Torr. ^{1}H-NMR (CDCl$_3$): 0.84–0.92 (m, 2 Me$_2$CHCH$_2$); 1.69 (m, 2 Me$_2$CHCH$_2$); 2.35 (m, 2 Me$_2$CHCH$_2$); 3.34 (s, 2 MeO); 4.47 (s, CH). ^{13}C-NMR (CDCl$_3$): 20.7; 27.0; 54.0; 56.8; 112.7; 112.8.

2′-Deoxy-N^2-[(diisobutylamino)methylidene]-5-methylisocytidine (**1h**). According to [18][23], **1d** (30.9 mmol, 7.45 g) was coevaporated with DMF and then suspended in dry DMF (150 ml). *N,N*-Diisobutylformamide dimethyl acetal (1.5 equiv., 46.4 mmol, 9.41 g) was added, and the suspension became a clear soln. within 10 min. It was stirred at r.t. for 1 h. Then MeOH (5 ml) was added and the mixture evaporated. The residue was purified by CC (silica gel, CHCl$_3$/MeOH 82.5:17.5; R_f 0.54): white solid (10.8 g, 92%). ^{1}H-NMR (CDCl$_3$): 0.90 (m, 2 Me$_2$CHCH$_2$); 1.95 (s, Me–C(5)); 1.98–2.39 (m, 1H–C(2′), 2 Me$_2$CHCH$_2$); 3.11 (m, 1H–C(2′)); 3.26–3.50 (m, 2 Me$_2$CHCH$_2$); 3.73–3.88 (m, 2H–C(5′)); 3.62 (m, H–C(4′)); 4.40 (m, H–C(3′)); 6.69 (dd, H–C(1′)); 8.04 (s, H–C(6)); 8.72 (s, N=CH). ^{13}C-NMR (CDCl$_3$): 14.0; 20.0; 20.6; 27.7; 28.2; 42.6; 54.8; 61.3; 62.6; 71.8; 88.2; 88.3; 89.1; 117.5; 137.0; 137.1; 159.2; 160.6; 175.1.

2′-Deoxy-N^2-[(diisobutylamino)methylidene]-5′-(4,4′-dimethoxytrityl)-5-methylisocytidine (**1i**). Compound **1h** (28.18 mmol, 10.71 g) was co-evaporated with pyridine and dissolved in dry pyridine (150 ml). DMAP (0.25 equiv., 7.05 mmol, 860 mg), (MeO)$_2$TrCl (1.15 equiv., 32.4 mmol, 11.0 g; 11.6 g, 95%), and Et$_3$N (2 equiv., 56.4 mmol, 7.84 ml) were added, and the mixture was stirred at r.t. for 2.5 h. MeOH (5 ml) was added and the

mixture evaporated. The residue was partitioned in aq. NaHCO$_3$ soln./AcOEt and extracted with AcOEt (3 × 200 ml). The combined org. layer was dried (Na$_2$SO$_4$) and evaporated and the residue purified by CC (silica gel, AcOEt/MeOH/Et$_3$N 80:2:1; R_f 0.21): yellow foam (9.97 g, 52%). ^{1}H-NMR (CDCl$_3$): 0.85 (m, 2 Me_2CHCH$_2$); 1.62 (s, Me–C(5)); 1.83–2.47 (m, 2H–C(2'), 2 Me$_2$CHCH$_2$); 2.98–3.55 (m, 2H–C(5'), 2 Me$_2$CHCH$_2$); 3.78 (2s, 2 MeO); 4.09 (m, H–C(4')); 4.58 (m, H–C(3')); 6.74 (dd, H–C(1')); 6.80 (m, 4 H, (MeO)$_2$Tr); 7.14–7.34 (m, 9 H, (MeO)$_2$Tr); 7.68 (s, H–C(6)); 8.54 (s, N=CH). ^{13}C-NMR (CDCl$_3$): 13.8; 19.7; 20.3; 26.4; 27.3; 41.9; 53.6; 55.1; 60.2; 63.5; 71.6; 85.9; 86.4; 86.7; 113.1; 117.7; 127.0; 127.9; 128.1; 129.1; 129.0; 133.4; 135.5; 136.0; 144.4; 157.3; 158.6; 159.0; 172.6.

2'-Deoxy-N^2-*[(diisobutylamino)methylidene]-5'-(4,4'-dimethoxytrityl)-5-methylisocytidine 3'-(2-Cyanoethyl Diisopropylphosphoramidite)* (**1j**). To a soln. of **1i** (13.4 mmol, 9.11 g) in CH$_2$Cl$_2$, *N,N*-diisopropylethylamine (5.5 equiv., 73.5 mmol, 12.8 ml) and 2-cyanoethyl diisopropylphosphoramidochloridite (1.1 equiv., 14.7 mmol, 3.28 ml) were added *via* syringe at 0°. The mixture was stirred at 0° for 5 min and at r.t. for 15 min. Aq. NaHCO$_3$ soln. (150 ml) was added and the mixture extracted with CH$_2$Cl$_2$/Et$_3$N 98:2. The combined org. layer was dried (Na$_2$SO$_4$) and evaporated and the residue purified by CC (silica gel, CH$_2$Cl$_2$/AcOEt/Et$_3$N 7:3:1; R_f 0.60). The obtained product was further purified by reprecipitation. A 50% soln. of **1j** in CH$_2$Cl$_2$/Et$_3$N 2:1 was added dropwise into vigorously stirred hexane causing the product to precipitate. Drying under vacuum gave 6.63 g (56%) of white foam. The hexane soln. was evaporated, the residue purified by CC and further purified by reprecipitation to give another 2.64 g (22.4%). Total yield: 9.27 g (79%). ^{1}H-NMR (CDCl$_3$): 0.82–0.96 (m), 1.04 (d), 1.13–1.20 (m), 1:22–1.29 (m, 24 H, Me_2CHCH$_2$, Me_2CH); 1.62 (s, Me–C(5)); 1.90–2.61 (m, 6 H, 2H–C(2'), CH$_2$CH$_2$CN, Me$_2$CHCH$_2$); 3.11–3.84 (m, 16 H, MeO (3.77), Me$_2$CH, Me$_2$CHCH$_2$, CH$_2$CH$_2$CN, 2H–C(5')); 4.20 (m, H–C(4')); 4.61 (m, H–C(3')); 6.74 (m, H–C(1')); 6.82 (m, 4 H, (MeO)$_2$Tr); 7.16–7.36, 7.40–7.46 (m, 9 H, (MeO)$_2$Tr); 7.69, 7.77 (2s, H–C(6)); 8.88 (s, N=CH). ^{13}C-NMR (CDCl$_3$): 13.4; 19.5; 19.9; 20.1; 24.2; 26.1; 27.0; 40.6; 40.8; 42.8; 43.0; 53.3; 54.9; 57.8; 58.0; 59.9; 62.7; 63.0; 72.7; 73.6; 85.0; 85.4; 86.1; 86.5; 112.9; 117.1; 117.7; 126.8; 127.7; 127.9; 127.9; 129.8; 133.0; 135.1; 135.2; 144.1; 157.1; 157.1; 158.4; 158.9; 172.3. ^{31}P-NMR (CDCl$_3$): 151.829; 152.591.

N^2-*Benzoyl-2'-deoxy-5'-O-(4,4'-dimethoxytrityl)-5-methylisocytidine 3'-O-(2-Cyanoethyl Phosphonate)* (**1l**). As described for similar reactions [15], **1c** (1 mmol, 847 mg) was dissolved in MeCN (5 ml), and 1*H*-tetrazole (0.5 mmol, 35 mg) in H$_2$O (0.1 ml) was added. The mixture was stirred at r.t. for 1.5 h and then evaporated. The residue was extracted (NaHCO$_3$ soln./AcOEt), the org. layer dried (Na$_2$SO$_4$) and evaporated, and the residue purified by reprecipitation from CH$_2$Cl$_2$ into hexanes (659 mg, 0.863, 86%). TLC (AcOEt/MeOH/Et$_3$N 80:2:1): R_f 0.55. ^{1}H- and ^{13}C-NMR: no i-Pr group. ^{31}P-NMR: no signals at 149 and 150 (*cf.* **1c**); instead, signals at 11.13 and 11.27 ppm.

N^2-*Benzoyl-2'-deoxy-5'-O-(4,4'-dimethoxytrityl)-5-methylisocytidine* (**1f**). Similarly to [16], **1l** (0.685 mmol, 523 mg) was treated with 10 ml of 2M KF in MeOH and stirred at r.t. for 5 h. The reaction was quenched with sulfur, the solvent evaporated, and the residue worked up by extraction with NaHCO$_3$ soln./AcOEt. The org. layer was dried (Na$_2$SO$_4$) and evaporated, and the residue purified by CC (CHCl$_3$/MeOH 98.5:1.5, R_f 0.30): light-yellow foam (249 mg, 56%). NMR: identical with that of **1f** obtained above.

2'-Deoxy-5'-(4,4'-dimethoxytrityl)-5-methylisocytidine (**1m**). Compound **1f** (0.167 mmol, 108 mg) was treated with hydrazine monohydrate (1 ml) in pyridine (1 ml). The two-layer mixture was vigorously stirred for 5¾ h, then H$_2$O was added and the mixture extracted with CHCl$_3$. The org. layer was dried (Na$_2$SO$_4$) and evaporated and the residue purified by CC (CHCl$_3$/MeOH, 82.5:17.5, R_f 0.34): white foam (49 mg, 54%). ^{1}H-NMR ((D$_6$)DMSO): 1.42 (s, Me–C(5)); 2.28 (m, 2H–C(2')); 3.20 (m, 2H–C(5')); 3.77 (s, 2 MeO); 3.88 (m, H–C(4')); 4.37 (m, H–C(3')); 5.40 (br., OH), 5.98 (dd, H–C(1')); 6.83 (m, 4 H, (MeO)$_2$Tr); 7.21–7.40 (m, 9 H, (MeO)$_2$Tr. ^{13}C-NMR ((D$_6$)DMSO): 13.2 (C(2') superimposed by DMSO); 55.0; 55.1; 63.3; 69.9; 85.5; 85.8; 113.2; 114.7; 126.8; 127.7; 127.9; 129.7; 132.4; 135.3; 135.4; 144.7; 154.2; 158.1; 168.6.

Nucleoside **1m** can be employed for formamidine protection and further for phosphoramidation.

2'-Deoxyisoguanosine (**2a**) was prepared from 2-amino-2'-deoxyadenosine (**7a**) in 81% yield using a slightly modified literature method [24].

2'-Deoxy-N^6-*[(diisobutylamino)methylidene]isoguanosine* (**2b**). Compound **2a** (0.807 mmol, 215.4 mg) was co-evaporated with DMF and dissolved in anh. DMF (5 ml). *N,N*-Diisobutylformamide dimethyl acetal (1.4 equiv., 1.58 mmol, 321 mg) was added and the mixture stirred at r.t. for 6.5 h. After evaporation, the residue was purified by CC (silica gel, CHCl$_3$/MeOH 82.5:17.5, R_f 0.35): off-white foam (315 mg, 96%). ^{1}H-NMR (CDCl$_3$): 0.67, 0.77 (2s, 2 Me_2CHCH$_2$); 1.89–2.21 (m, 2 Me$_2$CHCH$_2$); 2.40, 2.85 (2m, 2H–C(2')); 3.27, 3.45 (2m, 2 Me$_2$CHCH$_2$); 3.67–3.95 (m, 2H–C(5')); 4.17 (m, H–C(4')); 4.76 (m, H–C(3')); 6.30 (t, J = 6.9, H–C(1')); 7.90 (s, H–C(8)); 9.42 (s, N=CH). ^{13}C-NMR (CDCl$_3$): 19.5; 19.9; 26.8; 26.9; 39.8; 52.9; 60.3; 62.6; 71.7; 85.9; 88.6; 115.4; 140.9; 154.2; 156.1; 157.4; 162.4.

2′-Deoxy-N^6-*[(diisobutylamino)methylidene]*-3′,5′-*bis*-O-*(trimethylsilyl)isoguanosine* (**2c**). Compound **2b** (0.775 mmol, 315 mg) was co-evaporated with pyridine and dissolved in anh. THF (15 ml). Et$_3$N (3 equiv., 2.32 mmol, 0.323 ml) and Me$_3$SiCl (3 equiv., 2.32 mmol, 0.295 ml) were added, and the mixture was stirred at r.t. overnight (15 h). MeOH (2 ml) was added and the mixture evaporated. The residue was extracted (dil. NaHCO$_3$ soln./AcOEt) and the org. layer dried (Na$_2$SO$_4$) and evaporated. The product was pure by TLC (CHCl$_3$/MeOH 9:1, R_f 0.52) and the white foam (328 mg, 77%) used without further purification. ^{1}H-NMR (CDCl$_3$): 0.06 (*s*, 2 Me$_3$Si); 0.87 (*m*, 2 *Me*$_2$CHCH$_2$); 1.84–2.13 (*m*, 2 Me$_2$C*H*CH$_2$, 1H–C(2′)); 2.96 (*m*, 1H–C(2′)); 3.3 (*m*, Me$_2$CHC*H*$_2$); 3.40 (*m*, Me$_2$CHC*H*$_2$); 3.63 (*m*, 1H–C(5′)); 3.84 (*m*, 1H–C(5′)); 3.99 (*m*, H–C(4′)); 4.56 (*m*, H–C(3′)); 6.10 (*dd*, H–C(1′)); 7.55 (*s*, H–C(8)); 9.21 (*s*, N=CH). ^{13}C-NMR (CDCl$_3$): –0.2; 19.4; 19.8; 26.2; 26.8; 39.9; 52.6; 60.0; 62.8; 73.3; 86.8; 89.9; 115.9; 140.2; 154.3; 156.2; 157.1; 161.6.

2′-Deoxy-N^6-*[(diisobutylamino)methylidene]*-O^2-*(diphenylcarbamoyl)*-3′,5′-*bis*-O-*(trimethylsilyl)isoguanosine* (**7b**). Compound **2c** (0.596 mmol, 328 mg) was co-evaporated with pyridine and dissolved in anh. pyridine (10 ml). Diphenylcarbamoyl chloride (2 equiv., 1.19 mmol, 276 mg) and (i-Pr)$_2$EtN (1.7 equiv., 1.01 mmol, 0.176 ml) were added, and the mixture was stirred at r.t. for 4 h. MeOH (1 ml) was added, the mixture evaporated, the residue extracted (NaHCO$_3$/CHCl$_3$), the org. layer dried (Na$_2$SO$_4$) and evaporated, and the residue purified by CC (CH$_2$Cl$_2$/(MeOH 98.5:1.5; R_f 0.48): off-white foam (332 mg, 75%). ^{1}H-NMR (CDCl$_3$): 0.02 (*s*, 2 Me$_3$Si); 0.77 (*m*, 2 *Me*$_2$CHCH$_2$); 1.74–2.15 (*m*, 2 Me$_2$C*H*CH$_2$, 1H–C(2′)); 2.75 (*m*, 1H–C(2′)); 3.00 (*m*, Me$_2$CHC*H*$_2$); 3.41 (*m*, Me$_2$CHC*H*$_2$); 3.64–3.88 (*m*, 2H–C(5′)); 3.96 (*m*, H–C(4′)); 4.58 (*m*, H–C(3′)); 6.22 (*dd*, H–C(1′)); 7.14–7.50 (*m*, 10 H, Ph); 8.89 (*s*, N=CH). ^{13}C-NMR (CDCl$_3$): –0.6; 19.1; 19.5; 25.6; 26.5; 40.4; 52.0; 59.3; 62.0; 72.3; 75.8; 88.8; 125.0; 125.8; 126.3; 128.3; 140.8; 141.7; 151.2; 151.5; 155.0; 159.2; 161.0.

2′-Deoxy-N^6-*[(diisobutylamino)methylidene]*-O^2-*(diphenylcarbamoyl)isoguanosine* (**7c**). a) *From* **7b**. To a soln. of **7b** (0.445 mmol, 332 mg) in THF (10 ml), 1M Bu$_4$NF in THF (2.2 equiv., 0.98 mmol, 0.98 ml) was added. The mixture was stirred at r.t. for 1.5 h and then evaporated. The residue was extracted (H$_2$O/CHCl$_3$), the org. layer dried (Na$_2$SO$_4$) and evaporated, and the product purified by CC (CHCl$_3$/MeOH 9:1; R_f 0.38): off-white foam (238 mg, 89%).

b) *From* **9c**. To a soln. of **9c** (2.34 mmol, 1.97 g) in anh. THF (100 ml), 1M Bu$_4$NF in THF (2.2 equiv., 5.15 mmol, 5.15 ml) was added. The mixture was stirred at r.t. for 2 h and then evaporated. The residue was diluted with H$_2$O (50 ml) and extracted with CHCl$_3$. The org. layer was dried (Na$_2$SO$_4$) and evaporated, and the residue purified by CC (silica gel, CHCl$_3$/MeOH 9:1): light-yellow foam (1.25 g, 89%). ^{1}H-NMR (CDCl$_3$): 0.81–1.02 (*m*, 2 *Me*$_2$CHCH$_2$); 1.86–2.18 (*m*, 2 Me$_2$C*H*CH$_2$); 2.20 (*m*, 1H–C(2′)); 2.62 (*m*, 1H–C(2′)); 3.12 (*d*, Me$_2$CHC*H*$_2$); 3.43 (*m*, Me$_2$CHC*H*$_2$); 3.75–3.98 (*m*, 2H–C(5′)); 4.09 (*m*, H–C(4′)); 4.59 (*m*, H–C(3′)); 5.92 (*m*, H–C(1′)); 7.08–7.50 (*m*, 2 Ph); 7.74 (*s*, H–C(8)); 8.92 (*s*, N=CH). ^{13}C-NMR (CDCl$_3$): 19.6; 20.1; 26.0; 27.0; 40.7; 52.4; 59.8; 63.1; 72.6; 72.7; 87.3; 88.5; 125.4; 126.6; 126.6; 128.8; 141.8; 141.9; 150.9; 152.3; 154.8; 159.6; 161.3.

2′-Deoxy-N^6-*[(diisobutylamino)methylidene]*-5′-*(4,4′-dimethoxytrityl)*-O^2-*(diphenylcarbamoyl)isoguanosine* (**7d**). Compound **7c** (2.08 mmol, 1.25 g) was co-evaporated with pyridine, dissolved in anh. pyridine (50 ml) and DMAP (0.25 equiv., 0.521 mmol, 63.5 mg), (MeO)$_2$TrCl (1.8 equiv., 3.75 mmol; 1.34 g, 95%) and Et$_3$N (3 equiv., 6.25 mmol, 0.869 ml) were added. The mixture was stirred at r.t. for 4.5 h, then MeOH (5 ml) was added and the mixture evaporated. The residue was diluted with AcOEt and extracted (aq. NaHCO$_3$ soln./AcOEt), the combined org. layer dried (Na$_2$SO$_4$) and evaporated, and the residue purified by CC (silica gel, CHCl$_3$/MeOH 98.5:1.5, then 90:10; R_f 0.79): light-yellow foam (1.42 g, 76%). ^{1}H-NMR (CDCl$_3$): 0.90 (*dd*, 2 *Me*$_2$CHCH$_2$); 1.91–2.18 (*m*, 2 Me$_2$C*H*CH$_2$); 2.51 (*m*, 2H–C(2′)); 3.18 (*d*, Me$_2$CHC*H*$_2$); 3.30 (*m*, Me$_2$CHC*H*$_2$); 3.54 (*m*, 2H–C(5′)); 3.72 (*s*, 2 MeO); 4.09 (*m*, H–C(4′)); 4.54 (*m*, H–C(3′)); 6.47 (*dd*, H–C(1′)); 6.71–6.84 (*m*, 4 H, (MeO)$_2$*Tr*); 7.11–7.40 (*m*, 19 H, Ph, (MeO)$_2$Tr); 7.97 (*s*, H–C(8)); 8.97 (*s*, N=CH). ^{13}C-NMR (CDCl$_3$): 19.7; 20.0; 26.1; 27.0; 40.9; 52.6; 55.0; 55.0; 59.9; 63.8; 71.7; 71.8; 83.5; 83.6; 85.9; 86.3; 113.0; 124.4; 126.3; 126.7; 126.8; 127.7; 127.7; 128.0; 128.8; 129.0; 129.8; 129.9; 135.5; 135.6; 139.6; 142.1; 144.5; 152.2; 152.3; 156.0; 158.3; 158.4; 159.5; 161.2.

2′-Deoxy-N^6-*[(diisobutylamino)methylidene]*-5′-*(4,4′-dimethoxytrityl)*-O^2-*(diphenylcarbamoyl)isoguanosine* 3′-*(2-Cyanoethyl Diisopropylphosphoramidite)* (**7e**). To a soln. of **7d** (0.199 mmol, 180 mg) in CH$_2$Cl$_2$ (3 ml) at 0°, (i-Pr)$_2$EtN (5.5 equiv., 1.1 mmol, 0.19 ml) and 2-cyanoethyl diisopropylphosphoramidochloridite (1.1 equiv., 0.22 mmol, 0.049 ml) were added. The mixture was stirred at r.t. for 30 min. Then aq. NaHCO$_3$ soln. was added, the mixture extracted with CH$_2$Cl$_2$/Et$_3$N 98:2, the combined org. layer dried (Na$_2$SO$_4$) and evaporated, and the residue purified by CC (silica gel, hexanes/CHCl$_3$/AcOEt/Et$_3$N 60:20:20:10 (R_f 0.53, long spot), then 40:25:25:10 (R_f 0.36, 0.28 for the two diastereoisomers)). This product was obtained as a light-yellow foam. The obtained product was dissolved in CH$_2$Cl$_2$/10% Et$_3$N and added dropwise into stirred hexanes. The resulting precipitate was dried *in vacuo*: white foam (162 mg, 73.9%). ^{1}H-NMR (CDCl$_3$): 0.87–1.22 (*m*, 24 H, *Me*$_2$CHCH$_2$, *Me*$_2$CH); 1.90–2.23 (*m*, 2 Me$_2$C*H*CH$_2$); 2.37–2.71 (3*m*, 4 H, 2H–C(2′), CH$_2$CH$_2$CN), 3.18–3.79

(m, Me$_2$CHCH_2, Me$_2$CH, 2H–C(5'), CH$_2$CH$_2$CN, MeO (s, 3.78)); 4.23 (m, H–C(4')); 4.73 (m, H–C(3')); 6.46 (dd, H–C(1')); 6.75–6.84 (m, 4 H, (MeO)$_2$$Tr$); 7.14–7.46 ($m$, 19 H, Ph, (MeO)$_2$$Tr$); 8.02, 8.09 (2$s$, H–C(8)); 8.99 ($s$, N=CH). ^{13}C-NMR (CDCl$_3$): 19.7; 20.0; 24.4; 26.1; 27.1; 40.2; 43.0; 43.2; 52.5; 55.0; 55.1; 58.1; 58.3; 59.8; 63.1; 63.3; 73.4; 83.5; 85.5; 86.3; 113.0; 117.3; 117.5; 124.5; 124.6; 126.2; 126.7; 127.7; 127.9; 128.0; 128.8; 129.9; 130.0; 135.5; 135.5; 139.7; 142.2; 144.4; 152.1; 152.3; 152.4; 156.1; 156.1; 158.3; 159.5; 161.3. ^{31}P-NMR (CDCl$_3$): 151.497; 151.527.

Isoguanosine (**2g**). To a suspension of 2-aminoadenosine (**7f**; 18 mmol, 5.0 g) in H$_2$O (160 ml) at 50°, NaNO$_2$ (3.86 equiv., 68.4 mmol, 4.72 g) in H$_2$O (30 ml) was added. Then, AcOH (6.93 equiv., 123 mmol, 7.02 ml) was added dropwise at 50°. The resulting clear soln. was stirred for 5 min and then diluted with H$_2$O (150 ml). Conc. aq. ammonia was added to pH 8, the mixture evaporated, and the remaining solid washed with H$_2$O: light-yellow powder (4.59 g, 91.4%). ^{1}H-NMR ((D$_6$)DMSO): 3.50–3.74 (m, 2H–C(5')); 3.98 (m, H–C(4')); 4.12 (m, H–C(3')); 4.54 (m, H–C(2')); 5.69 (m, H–C(1')); 7.97 (s, H–C(8)). ^{13}C-NMR ((D$_6$)DMSO): 61.9; 71.0; 73.2; 86.3; 87.9; 109.8; 138.38; 152.78; 153.58; 156.8.

N^6-*[(Diisobutylamino)methylidene]isoguanosine* (**2h**). Compound **2g** (16.2 mmol, 4.59 g) was co-evaporated with pyridine followed by DMF and suspended in dry DMF (100 ml). *N,N*-Diisobutylformamide dimethyl acetal (1.4 equiv., 22.7 mmol, 4.61 g) was added, the mixture stirred at r.t. for 23 h, then MeOH (3 ml) added, and the mixture evaporated. The residue was purified by CC (silica gel, CHCl$_3$/MeOH 82.5:17.5; R_f 0.38): light-yellow foam (6.17 g, 90%). ^{1}H-NMR ((D$_6$)DMSO): 0.81–0.98 (m, 2 Me_2CHCH$_2$); 1.97–2.19 (m, 2 Me$_2$CHCH$_2$); 3.28–3.50 (m, 2 Me$_2$CHCH_2); 3.60–3.78 (m, 2H–C(5')); 4.07 (m, H–C(4')); 4.25 (m, H–C(3')); 4.59 (m, 1H–C(2')); 5.79 (m, H–C(1')); 8.18 (s, H–C(8)); 9.21 (s, N=CH). ^{13}C-NMR ((D$_6$)DMSO): 20.2; 20.6; 26.8; 27.5; 53.2; 60.6; 62.3; 71.3; 74.1; 86.6; 88.5; 114.7; 141.9; 155.6; 157.4; 157.7; 163.1.

N^6-*[(Diisobutylamino)methylidene]-3',5'-O-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)isoguanosine* (**8**). Compound **2d** (0.761 mmol, 321 mg) was co-evaporated with pyridine and dissolved in dry pyridine (5 ml), 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane (1.1 equiv., 0.837 mmol, 0.268 ml) added, and the mixture stirred at r.t. for 6 h. Aq. NaHCO$_3$ soln. (10 ml) was added and the mixture extracted (CHCl$_3$). The combined org. layer was dried (Na$_2$SO$_4$) and evaporated and the residue used for NMR analysis: mixture of desired **8** and the corresponding 5'-O-(3-hydroxy-1,1,3,3-tetraisopropyldisiloxan-1-yl) compound **2e**. Separation was possible but the yield was too poor for large-scale use.

Alternatively, according to [25], **2h** (0.3507 mmol, 148 mg) was co-evaporated with pyridine and dissolved in dry DMF (5 ml), 1*H*-imidazole (4 equiv., 1.40 mmol, 95.4 mg) and 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane (1.1 equiv., 0.386 mmol, 0.123 ml) were added, and the mixture was stirred at r.t. until the reaction was complete (30 min). Aq. NaHCO$_3$ soln. (10 ml) was added and the mixture extracted (CHCl$_3$). The combined org. layer was dried (Na$_2$SO$_4$), and evaporated and the residue used for NMR analysis: **8/2e**.

N^6-*[(Diisobutylamino)methylidene]-2',3',5'-tris-O-(trimethylsilyl)isoguanosine* (**2i**). Compound **2h** (8.23 mmol, 3.48 g) was co-evaporated with pyridine and dissolved in anh. THF (100 ml). Et$_3$N (4 equiv., 32.9 mmol, 4.58 ml) and Me$_3$SiCl (4 equiv., 32.9 mmol, 4.18 ml) were added, and a white precipitate was formed immediately. The mixture was stirred at r.t. overnight (13 h). TLC (CHCl$_3$/MeOH 9:1): R_f 0.48 (**2i**); 0.27 (starting material). MeOH (2 ml) was added, the mixture evaporated, dil. aq. NaHCO$_3$ soln. added, and the mixture extracted with AcOEt. The org. layer was dried (Na$_2$SO$_4$), the solvent evaporated, and the obtained light-yellow foam (5.04 g, 96%) used in the next step without further purification. ^{1}H-NMR (CDCl$_3$): 0.10, 0.14, 0.18 (3s, 3 Me$_3$Si); 0.91–1.10 (m, 2 Me_2CHCH$_2$); 2.03–2.32 (m, 2 Me$_2$CHCH$_2$); 3.32–3.61 (m, 2 Me$_2$CHCH_2); 3.76–4.04 (m, 2H–C(5')); 4.16 (m, H–C(4')); 4.34 (m, H–C(3')); 4.58 (m, H–C(2')); 6.07 (m, H–C(1')); 8.04 (s, H–C(8)); 9.41 (s, N=CH). ^{13}C-NMR (CDCl$_3$): –0.8; 0.1; 19.7; 20.0; 26.4; 27.1; 52.8; 60.7; 61.0; 70.9; 75.5; 83.9; 87.6; 114.6; 139.5; 153.7; 157.7; 158.0; 161.9.

N^6-*[(Diisobutylamino)methylidene]-O^2-(diphenylcarbamoyl)-2',3',5'-tris-O-(trimethylsilyl)isoguanosine* (**7g**). Compound **2i** (7.91 mmol, 5.04 g) was co-evaporated with pyridine and dissolved in anh. pyridine (100 ml). (i-Pr)$_2$EtN (1.7 equiv., 13.4 mmol, 2.34 ml) and diphenylcarbamoyl chloride (2 equiv., 15.8 mmol, 3.66 g) were added. The dark-orange soln. was stirred at r.t. for 2.25 h, then dil. aq. NaHCO$_3$ soln. added with ice-bath cooling, and the mixture extracted with CHCl$_3$. The org. layer was dried (Na$_2$SO$_4$) and evaporated and the residue purified by CC (CH$_2$Cl$_2$/MeOH 98.5:1.5): yellow foam (5.11 g, 78%). ^{1}H-NMR (CDCl$_3$): 0.10, 0.14, 0.18 (3s, 3 Me$_3$Si); 0.93 (dd, 2 Me_2CHCH$_2$); 1.93–2.20 (m, 2 Me$_2$CHCH$_2$); 3.18 (d, Me$_2$CHCH_2); 3.47–3.64 (m, Me$_2$CHCH_2); 3.70–4.05 (2m, 2H–C(5')); 4.12 (m, H–C(4')); 4.31 (m, H–C(3')); 4.51 (m, H–C(2')); 6.00 (m, H–C(1')); 7.22–7.40 (m, 2 Ph); 8.27 (s, H–C(8)); 8.94 (s, N=CH). ^{13}C-NMR (CDCl$_3$): –0.7; 0.1; 19.8; 20.1; 26.1; 27.2; 52.7; 59.9; 60.8; 70.4; 75.8; 83.8; 88.9; 123.7; 126.2; 126.9; 128.8; 140.5; 140.5; 142.3; 152.0; 152.4; 156.0; 159.4; 161.2.

N^6-*[(Diisobutylamino)methylidene]-O^2-(diphenylcarbamoyl)isoguanosine* (**7h**). Compound **7g** (6.13 mmol, 5.11 g) was dissolved in anh. THF (100 ml) and 1M Bu$_4$NF in THF (3.3 equiv., 20.2 mmol, 20.2 ml) added *via*

syringe. The mixture was stirred at r.t. for 1.5 h and then evaporated. H_2O was added and the mixture extracted with $CHCl_3$. The org. layer was dried (Na_2SO_4) and evaporated and the residue purified by CC (silica gel, $CHCl_3/MeOH$ 9:1; R_f 0.27): light-yellow foam (3.44 g, 91%). ^1H-NMR ($CDCl_3$): 0.88 (m, 2 Me_2CHCH_2); 1.83–2.17 (m, 2 Me_2CHCH_2); 3.18 (m, Me_2CHCH_2); 3.37–3.58 (m, Me_2CHCH_2); 3.63–3.80 (2m, 2H–C(5')); 4.16 (m, H–C(4')); 4.29 (m, H–C(3')); 4.82 (m, H–C(2')); 5.37 (m, H–C(1')); 7.18–7.48 (m, 2 Ph); 7.63 (s, H–C(8)); 8.99 (s, N=CH). ^{13}C-NMR ($CDCl_3$): 19.6; 20.1; 26.4; 27.0; 52.6; 60.0; 63.1; 72.4; 73.6; 87.1; 90.2; 125.4; 126.7; 127.1; 128.9; 141.8; 142.3; 151.0; 152.8; 154.8; 160.2; 161.4.

N^6-[(Diisobutylamino)methylidene]-O^2-(diphenylcarbamoyl)-3',5'-O-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)isoguanosine (9a). Compound 7h (5.29 mmol, 3.27 g) was co-evaporated with pyridine and dissolved in anh. pyridine (70 ml), 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane (1.1 equiv., 5.82 mmol, 1.86 ml) added via syringe, and the mixture stirred for 16.25 h at r.t. The mixture was evaporated, dil. $NaHCO_3$ soln. added. The mixture extracted with $CHCl_3$, the org. layer dried (Na_2SO_4) and evaporated, and the residue purified by CC (silica gel, $CHCl_3/MeOH$ 9:1; R_f 0.77): light-yellow foam (4.18 g, 92%). ^1H-NMR ($CDCl_3$): 0.90–1.19 (m, 40 H, Me_2CH, Me_2CHCH_2); 1.93–2.20 (m, 2 Me_2CHCH_2); 3.19 (d, Me_2CHCH_2); 3.57 (m, Me_2CHCH_2); 4.03–4.12 (m, 2H–C(5'), H–C(4')); 4.49 (m, H–C(3')); 4.78 (m, H–C(2')); 6.00 (m, H–C(1')); 7.18–7.40 (m, 2 Ph); 7.99 (s, H–C(8)); 8.98 (s, N=CH). ^{13}C-NMR ($CDCl_3$): 12.5; 17.3; 19.8; 20.1; 26.2; 27.1; 52.7; 60.0; 61.9; 71.0; 75.1; 82.0; 88.9; 89.0; 123.7; 125.0; 126.2; 128.9; 140.0; 140.1; 142.3; 149.8; 152.0; 156.2; 159.6; 161.4.

N^6-[(Diisobutylamino)methylidene]-O^2-(diphenylcarbamoyl)-2'-O-(phenoxythiocarbonyl)-3',5'-O-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)isoguanosine (9b). Compound 9a (3.42 mmol, 2.94 g) was co-evaporated with pyridine and dissolved in CH_2Cl_2 (120 ml). Pyridine (4 equiv., 13.7 mmol, 1.11 ml), DMAP (0.25 equiv., 0.855 mmol, 104 mg), and phenyl carbonochloridathioate (1.5 equiv., 5.1 mmol, 0.71 ml) were added. The mixture turned brown-red, was stirred at r.t. for 19.5 h, and then extracted ($CHCl_3$/dil. $NaHCO_3$ soln.). The org. layer was dried (Na_2SO_4) and evaporated and the residue purified by CC (one-step gradient of $CH_2Cl_2/MeOH$ 82.5:1.5; R_f 0.44, then $CHCl_3/MeOH$ 9:1): light-yellow foam (3.21 g, 94%). ^1H-NMR ($CDCl_3$): 0.87–1.12 (m, 40 H, Me_2CH, Me_2CHCH_2); 1.92–2.20 (m, 2 Me_2CHCH_2); 3.18 (d, Me_2CHCH_2); 3.54 (m, Me_2CHCH_2); 4.00–4.28 (m, 2H–C(5'), H–C(4')); 4.95 (m, H–C(3')); 6.26 (m, H–C(1'), H–C(2')); 7.06–7.45 (m, 3 Ph); 8.06 (s, H–C(8)); 8.96 (s, N=CH). ^{13}C-NMR ($CDCl_3$): 12.7; 12.8; 12.9; 13.2; 17.0; 17.2; 19.8; 20.0; 26.2; 27.1; 52.7; 60.0; 60.7; 69.7; 82.1; 83.9; 86.4; 115.5; 121.4; 121.7; 124.8; 126.5; 126.9; 128.9; 129.4; 139.5; 142.2; 152.0; 153.4; 156.4; 159.5; 161.4; 193.6.

2'-Deoxy-N^6-[(diisobutylamino)methylidene]-O^2-(diphenylcarbamoyl)-3',5'-O-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)isoguanosine (9c). Compound 9b (3.22 mmol, 3.21 g) was co-evaporated with toluene, dissolved in anhy. toluene (75 ml), and degassed with Ar for 30 min. In a second flask, 2,2'-azobis[isobutyronitrile] (AIBN; 0.5 equiv., 1.61 mmol, 264 mg) and Bu_3SnH (3.0 equiv., 9.65 mmol, 2.60 ml) in toluene (15 ml) were degassed with Ar for 30 min. The first flask was then heated to 80°, the AIBN/Bu_3SnH soln. added dropwise via syringe, and the mixture stirred at 100° for 3.5 h; gas evolution ceased after 20 min. After cooling to r.t., the mixture was evaporated, the residue extracted with dil. $NaHCO_3$ soln./$CHCl_3$, the org. layer dried (Na_2SO_4) and evaporated, and the residue purified by CC (silica gel, $CHCl_3/MeOH$ 9:1; R_f 0.80): yellow oil (1.97 g, 73%). ^1H-NMR ($CDCl_3$): 0.78–1.11 (m, 40 H, Me_2CH, Me_2CHCH_2); 1.82–2.21, 2.52–2.66 (m, 2H–C(2'), 2 Me_2CHCH_2); 3.02–3.58 (d, 2 Me_2CHCH_2); 3.71–4.08 (m, H–C(4'), 2H–C(5')); 4.72 (m,H–C(3')); 6.29 (m, H–C(1')); 7.04–7.56 (m, 2 Ph); 8.16 (s, H–C(8)); 8.90 (s, N=CH). ^{13}C-NMR ($CDCl_3$): 12.4; 12.7; 12.9; 13.2; 16.9; 17.0; 17.2; 17.4; 19.6; 19.7; 27.0; 40.1; 52.7; 59.8; 61.5; 69.5; 82.9; 85.0; 119.4; 125.8; 126.5; 127.1; 128.8; 139.7; 142.0; 151.4; 152.2; 156.1; 159.3; 161.1.

Ion-Exchange HPLC. Oligonucleotides were chromatographed by anion-exchange HPLC with an Alltech Associates column (Macrosphere 300A WAx, 7 µm, 4.6 × 250 mm), using either of the two gradients with $A = 25$ mM Tris · HCl (pH 7.5/MeCN 9:1 and $B = 25$ mM Tris · HCl (pH 7.5) 2M NaCl/MeCN 9:1 or $C = 20$ mM K_3PO_4 (pH 6.0) and $D = 20$ mM K_3PO_4 (pH 6.0), 2M KCl; flow rate 0.5 ml/min. For anal. purposes, 0.1 A_{260} units were injected. Semi-prep. separations were done with the 4.6-mm diameter column as well, by injecting up to 6 A_{260} units.

Anal. Reversed-Phase HPLC. Oligonucleotides were analyzed by reversed-phase HPLC with a Waters Corporation column (Nova-Pak C_{18} 60A, 4 µm, 3.9 × 150 mm), using a gradient with $A = 100$ mM (Et$_3$NH)OAc (pH 7.0) and $B = $ MeCN; flow rate 0.5 ml/min; 0.1 A_{260} units were injected.

Semiprep. Reversed-Phase HPLC. After semiprep. ion-exchange HPLC with oligonucleotides, the solvent of collected fractions was lyophilized. The solns. were then further purified with a Waters Corporation column (Nova-Pak HRC$_{18}$ 60A, 6 µm, 7.8 × 300 mm), using a gradient with $A = 25$ mM (Et$_3$NH)OAc (pH 7.0) and $B = $ MeCN; flow rate 2.2 ml/min.

Reversed-Phase HPLC for Analysis of Nucleosides. Following digestion of oligonucleotides, the component nucleosides were separated by reversed-phase HPLC with a *Waters Corporation* column (*Nova-Pak C_{18} 60A*, 4 µm, 3.9 × 150 mm), using a gradient with A = 50 mM K_3PO_4 (pH 7.0) and B = 50 mM K_3PO_4 (pH 7.0)/MeOH 2:1; flow rate 0.5 ml/min. Order of eluting of 2'-deoxynucleosides: C_d, isoC_d, isoG_d, G_d, T_d, A_d.

Oligonucleotide Synthesis. Bottles to contain non-standard base phosphoramidites for the DNA synthesizer were washed, rinsed with acetone, and heated overnight at 120°. The bottles were then allowed to cool overnight in a dessicator over P_2O_5 and under vacuum of < 1 Torr. IsoC$_d$ phosphoramidite was introduced and the bottles containing phosphoramidite were again placed in the dessicator under vacuum overnight.

Oligonucleotide synthesis was performed by the DNA Synthesis Core of the University of Florida on a *PerSeptive Biosystems Expedite 8900* synthesizer, at the 0.2-µmol scale. The synthesis was done using the manufacturer's recommended procedure, except the concentrations of standard phosphoramidites were half the manufacturer's recommended concentration. All chemicals were from *PerSeptive Biosystems.* DNA Membrane 0.2 µm columns, also from *PerSeptive*, containing the 3'-terminal base of the desired oligonucleotide, were used (part numbers for each base: C = GEN050014, G = GEN050024, T = GEN050026, A = GEN050004). Anh. MeCN from *Glen Research* (part number 40-4050-45) was used to dilute the phosphoramidites (0.5 g in 10 ml for standard bases, 0.15 g in 1.5 ml for non-standard bases).

The synthesis protocols used are below. isoC$_d$ Phosphoramidite was coupled using the standard protocol, and isoG$_d$ phosphoramidite was coupled with an extended (600 s) coupling time.

At the completion of the synthesis, oligonucleotides were deprotected by incubating in 30% aq. ammonium hydroxide (1.5 ml) at 55° for 12 h. The sample was then centrifuged and the supernatant removed and dried to a pellet. After drying, 0.7 mM NaOAc (0.3 ml) was added, the soln. vortexed, and EtOH added (1 ml) at −20°. This soln. was stored at −20° for 15 min before being centrifuged (15000 g, 20 min). The supernatant was discarded and the pellet suspended in 80% EtOH/H_2O (1 ml) and centrifuged again (15000 g, 20 min). The desalted oligonucleotide was then redissolved in 10 mM *Tris* · HCl, 0.4 mM EDTA, pH 7.7 (0.5 ml).

Coupling Efficiency by (MeO)$_2$Tr Monitoring. A short oligomer having the sequence 5'-d(TTTTT-XXXXXTTTTT)-3' (X = non-standard base) was synthesized at 0.2-µmol scale using the phosphoramidite to be evaluated. (MeO)$_2$Tr Release at each coupling was monitored using the trityl viewer on an *Expedite 8909* synthesizer. The coupling efficiency of each of the three sections (T_5, X_5, and T_5) was then determined by using $100 \cdot 2^m$, where m is the slope of the linear regression line of natural logarithm of the absorbance value *vs.* the coupling number.

Enzymatic Digestion of Oligonucleotides. The enzymes used were alkaline phosphatase from bovine calf intestine (18.2 prot mg/ml in 50% glycerol, 5 mM $MgCl_2$, and 0.1 mM $ZnCl_2$, pH 7.5) and phosphordiesterase I from *Crotalus durissus terrificus* venom (0.16 mg/ml in 50% glycerol, 5 mM *Tris* · HCl, pH 7.5). A stock soln. of alkaline phosphatase was prepared by adding enzyme soln. (4 µl) to a soln. of 50 mM *Tris* · HCl, 10 mM $MgCl_2$, pH 8.3 (198 µl).

The oligonucleotide (0.1 *OD*) was dissolved in 0.1M *Tris* · HCl, 20 mM $MgCl_2$, pH 8.3 (2 µl). H_2O (2 µl), phosphodiesterase I (0.5 µl), and phosphatase stock soln. (0.5 µl) were added. The resulting soln. was incubated at 37° for 3 h, diluted with 1M (Et$_3$NH)OAc, (pH 7; 80 µl), filtered, and analyzed using reversed-phase HPLC. Composition of the oligonucleotide was verified by using the integrated absorbance of the component nucleosides with the following extinction coefficients (M^{-1} cm^{-1}): 15400 (A_d), 11700 (G_d), 7300 (C_d), 8800 (T_d), 10100 (U_d), 6300 (isoC$_d$), and 4600 (isoG$_d$).

REFERENCES

[1] J. A. Piccirilli, T. Krauch, S. E. Moroney, S. A. Benner, *Nature (London)* **1990**, *343*, 33.
[2] S. A. Benner, R. K. Alleman, A. D. Ellington, L. Ge, A. Glasfeld, G. F. Leanz, T. Krauch, L. J. MacPherson, S. E. Moroney, A. J. Piccirilli, E. Weinhold, *Cold Harbor Symp Quant. Biol.* **1987**, *52*, 53.
[3] A. Rich, in 'Horizons in Biochemistry', Eds. M. Kasha and B. Pullman, Academic Press, New York, 1962, pp. 103–126.
[4] Y. Tor, P. B. Dervan, *J. Am. Chem. Soc.* **1993**, *115*, 4461.
[5] T. Horn, C.-A. Chang, M. L. Collins, *Tetrahedron Lett.* **1995**, *36*, 2033.
[6] F. Seela, C. Wei, *Helv. Chim. Acta* **1997**, *80*, 73.
[7] C. Y. Switzer, S. E. Moroney, S. A. Benner, *Biochemistry* **1993**, *32*, 10489.
[8] C. Roberts, R. Bandrau, C. Switzer, *J. Am. Chem. Soc.* **1997**, *119*, 4640.
[9] C. Roberts, R. Bandrau, C. Switzer, *Tetrahedron Lett.* **1995**, *36*, 3601.

[10] H. Sugiyama, S. Ikeda, I. Saito, *J. Am. Chem. Soc.* **1996**, *118*, 9994.

[11] S. A. Strobel, T. R. Cech, N. Usman, L. Beigelman, *Biochemistry* **1994**, *33*, 13824.

[12] M. L. Collins, B. Irvine, D. Tyner, E. Fine, C. Zayati, C. Chang, T. Horn, D. Ahle, J. Detmer, L.-P. Shen, J. Kohlberg, S. Bushnell, M. S. Urdea, D. D. Ho, *Nucleic Acids Res.* **1997**, *25*, 2979.

[13] M. D. Matteucci, M. H. Caruthers, *Tetrahedron Lett.* **1980**, *21*, 3243.

[14] J. R. Hwu, M. L. Jain, S.-C. Tsay, G. H. Hakimelahi, *J. Chem. Soc., Chem. Common.* **1996**, 454.

[15] M. M. P. Ng, F. Benseler, T. Tuschl, F. Eckstein, *Biochemistry* **1994**, *33*, 12119.

[16] W. K.-D. Brill, *Tetrahedron Lett.* **1994**, *35*, 3041.

[17] F. Seela, C. Wei, Z. Kazimierczuk, *Helv. Chim. Acta* **1995**, *78*, 1843.

[18] D. Wang, P. O. P. Ts'o, *Nucleosides/Nucleotides* **1996**, *5*, 387.

[19] G. H. Hakimelahi, Z. A. Proba, K. K. Ogilvie, *Can. J. Chem.* **1982**, *60*, 1106.

[20] M. J. Damha, K. K. Ogilvie, in 'Protocols for Oligonucleotides and Analogs', Ed. S. Agrawal, Human Press, Totowa, N.J., 1993, Vol. 20, pp. 81–114.

[21] K. A. Watanabe, U. Reichman, C. K. Chu, J. J. Fox, in 'Nucleic Acid Chemistry', Eds. R. S. Tipson and C. B. Townshed, John Wiley and Sons, New York, 1978, Vol. 1, pp. 273–277.

[22] B. C. Froehler, M. D. Matteucci, *Nucleic Acids Res.* **1983**, *11*, 8031.

[23] L. L. McBridge, R. Kierzek, S. L. Beaucage, M. H. Caruthers, *J. Am. Chem. Soc.* **1986**, *108*, 2040.

[24] F. Seela, B. Gabler, *Helv. Chim. Acta* **1994**, *77*, 622.

[25] W. Markiewicz, *J. Chem. Res. (S)* **1979**, 24.

An *In Vitro* Screening Technique for Polymerases that can Incorporate Modified Nucleotides. Pseudouridine as a Substrate for Thermostable Polymerases

Stefan Lutz, Petra Burgstaller[§], Steven A. Benner*

Departments of Chemistry and Anatomy and Cell Biology
University of Florida, Gainesville, Florida 32611

[§] Present address: Aventis GmbH&Co KG, Frankfurt am Main, Germany

Abstract:

Structurally and functionally modified nucleotides expand the scope of therapeutic and diagnostic applications of oligonucleotides. Incorporation of these into DNA and replication of oligonucleotides containing them by polymerases may cause pausing, early termination, and misincorporation, however. Reported here is a simple *in vitro* assay to screen for DNA polymerases that accept modified nucleotides, based on a set of primer extension reactions which, in combination with the scintillation proximity assay (SPA™), allows rapid and simple screening of multiple enzymes for their accuracy of replicating oligonucleotides in the presence of unnatural nucleotides. A proof of concept is obtained using pseudo-thymidine (ψT; **1**), the C-nucleoside analog of thymidine, as an unnatural substrate. The conformational properties of **1** arising from the carbon-carbon bond between the sugar and the base make it an interesting probe for the importance of conformational restraints in the active site of polymerases during primer elongation. From a pool of commercially available thermostable polymerases, the assay identified *Taq* DNA polymerase as the most suitable enzyme for the PCR amplification of oligonucleotides containing **1**. Subsequent experiments analyzing processivity and fidelity of *Taq* DNA polymerase acting on **1** are presented.

1

Introduction:

The expansion of the nucleic acid-based chemistry and biology by novel, functionalized nucleotides and non-standard base pairs is driven by the possibility to identify RNA and DNA molecules with new catalytic and binding properties [Tarasow, 1997 #939][Sakthivel, 1998 #1346][Piccirilli, 1990 #44] and to better understand the physical organic chemistry underlying strand-strand interaction in nucleic acids (xxx Benner, SA, Battersby, TR, Kodra, J, Switzer, CY, Moroney, SE, Vögel, J, MacPherson, L, von Krosigk, U, Hammer, C, Richert, C, Huang, Z, Horlacher, J, Schneider, KC, König, M, Blättler, M, Arslan, T, Hyrup, B, Egli, M, Bäschlin, D, Müller, E, Schmidt, J, Piccirilli, J, Roughton, A, Held, H, Lutz, M, Eschgfäller, B, Jurczyk, S, Lutz, S, Hutter, D, Nambiar, KP, Stackhouse, J (1998) Redesigning nucleic acids. *Pure Appl. Chem.* **70**, 263-266.). Additional nucleotide building blocks in a DNA strand could also introduce novel structural motifs and functional groups, including thiols[Chaudhuri, 1995 #1428] and imidazoles[Kodra, 1997 #996] not found in standard nucleic acids. Sophisticated techniques such as *in vitro* selection [xxx Szostak, JW (1992) In vitro genetics. *Trends Biochem. Sci.* **17**, 89-93.] could be applied to explore the lengthened dimensions in sequence space to generate more powerful ligands, receptors, and catalysts than can be obtained from standard *in vitro* selection experiments.

To implement this vision requires molecular biological tools for manipulating an expanded genetic alphabet, however, tools having the same power with the nucleic acid derivatives as natural tools have for manipulating natural nucleic acids. Enzymes, in particular DNA polymerases, that utilize unnatural substrates with the same efficiency as the natural polymerases would be especially valuable. While many enzymes accept components of an expanded genetic alphabet (xxx Bain, JD, Chamberlin, AR, Switzer, CY, Benner, SA (1992) Ribosome-mediated incorporation of non-standard amino acids into a peptide through expansion of the genetic code. *Nature* **356**, 537-539.), natural polymerases have clearly been optimized for the structures of the natural nucleotides (xxx Horlacher, J, Hottiger, M, Podust, VN, Hübscher, U, Benner, SA (1995) Expanding the genetic alphabet: Recognition by viral and cellular DNA polymerases of nucleosides bearing bases with non-standard hydrogen bonding patterns. *Proc. Natl. Acad. Sci.* **92**, 6329-6333). When they encounter

2

an unnatural base either in the tempalte or as a triphosphate, they frequently pause or abort copying. Even when the pausing is modest, it can have consequences in a PCR amplification, where copies must be made of copies, and pausing ocurs at each round of copying.

For these reasons, polymerases optimized to handle an expanded genetic alphabet need to be found by experiment. To date, polymerases have been screened using gel-based assays that allow examination of tens to hundreds of variants. These experiments have detected polymerases with improved ability to accept non-standard bases (need reference xxx), and a crude statistical evaluation of the results suggests that satisfactory polymerases will be attained, not after millions of variants are inspected, but after thousands are inspected. The goal of a polymerase-based molecular biology of an expanded genetic alphabet is therefore accessible, but only if a higher throughput screen is developed. In view of the interest of now many laboratories in improving polymerases for a variety of nucleic acid analogs (Moran, S, Ren, RXF, Rumney, S, Kool, ET (1997) Difluorotoluene, a nonpolar isostere for thymine, codes specifically and efficiently for adenine in DNA replication. *J. Am. Chem. Soc.* **119**, 2056-2057;

Tarasow TM, Eaton BE (1998) Dressed for success. Realizing the catalytic potential of RNA. *Biopolymers* **48**, 29-37.

Seela, F, Wei, CF (1997) Oligonucleotides containing consecutive 2'-deoxyisoguanosine residues: Synthesis, duplexes with parallel chain orientation, and aggregation *Helv. Chim. Acta* **80**, 73-85.), it is appropriate to describe this screening method at this time.

Material and Methods

Chemical synthesis of pseudo-thymidine ψT (1) and ψTTP (3)

The phosphoramidite of pseudo-thymidine (ψT, **1**) was prepared from pseudouridine [Bhattacharya, 1995 #1124]. The triphosphate **3** (ψTTP) of **1** was synthesized from 3'-acetyl-pseudothymidine, using the procedure by Ludwig and Eckstein [Ludwig, 1989 #47]. The 5' position of pseudothymidine was protected with 4,4'-dimethoxytrityl chloride (in pyridine, 16 h, rt), followed by acetylation of the 3'-hydroxy group with acetic anhydride (4 moleq. in pyridine, 4 h, rt). After evaporation of the solvents, the residue was detritylated with trifluoroacetic acid in dichloromethane,

3

yielding 3'-acetyl-pseudothymidine. This intermediate (8.5 mg, 30 µmol) was dissolved in a mixture of pyridine and dioxane (30 µl / 89 µl) and reacted with a solution of 2-chloro-4H-1,3,2-benzodioxaphosphorin-4-one (36 µl; 188 mg in 0.93ml dioxane). After 10 minutes, tributylammonium pyrophosphate (17 mg; 37 µmol) in DMF (91 µl) and tributylamine (31 µl) was added and the reaction mixture shaken for another 10 minutes. The reaction was quenched by addition of iodine solution (0.6 ml; 1% in pyridine/water 98:2). After shaking the mixture at room temperature for 15 minutes, the excess iodine was destroyed with a few drops of Na_2SO_3 solution (5%). The solution was evaporated and dried completely under high vacuum. The residue was reacted with water (0.8 ml) at room temperature for 30 minutes. The solution was lyophilized and the residue incubated in ammonia (1 ml; conc.) at 55°C overnight. After degasing and lyophilizing the solution, the raw product was purified by reverse-phase HPLC (Waters Prep.Nova Pak HR C18, 7.8 x 300mm; 4 ml/min.; 10 mM triethylammonium acetate (pH 7.0) – 10 mM triethylammonium acetate (pH 7.0) + 20% acetonitrile; 0 to 25 % B in 30 minutes). The product fractions were pooled and lyophilized three times to remove excess triethyammonium acetate, yielding 2 mg (14%) of **3** as a white-yellow foam.

H-NMR (300MHz, D₂O) δ: 7.85 (s, 1H, H-6), 5.18 (q, 1H, H-1', J = 6.04, 9.61 Hz), 4.62 (hept, 1H, H-3', J = 2.5, 5.3 Hz), 4.20 (m, 1H, H-4'), 3.71 (dd, 1H, H-5'$_A$, J = 4.4, 11.7 Hz), 3.62 (dd, 1H, H-5'$_B$, J = 6.3, 11.7 Hz), 3.47 (s, 3H, H-(NCH₃), 3.27 (q, triethylammonium), 2.34 (ddd, 1H, H-2'$_A$, J = 2.47, 6.05, 13.5 Hz), 2.18 (ddd, 1H, H-2'$_B$, J = 5.7, 9.61, 13.5 Hz), 1.35 (t, triethylammonium), ppm. **P-NMR (300MHz, D₂O)** δ: 3.05 (d, γP, J = 18.5 Hz), 2.23 (d, αP, J = 24.4 Hz), -9.73 (t, βP, J = 18.5, 24.4 Hz), ppm. **MS:** (ESI⁻): 481 (M-H), 503 (M-H, Na). **UV:** (Tris-HCl, 100 mM, pH 7.0): $\lambda_{max.}$ = 271 nm (ϵ = 8800).

DNA synthesis, digestion, and sequencing

Oligonucleotides containing **1** were prepared on an Applied Biosystems DNA synthesizer. The coupling times for **1** were doubled. No further adjustments of the protocol were necessary. The oligonucleotides were deprotected under standard conditions (ammonia solution, 55°C, overnight) and PAGE-purified[Sambrook, 1989 #341].

4

Successful incorporation of **1** by chemical synthesis or enzymatic primer extension was shown through enzymatic digestion of the oligonucleotides with phosphodiesterase (*Crotalus durissus terrificus*, Boehringer Mannheim) and alkaline phosphatase (*Bovine calf intestine*, Boehringer Mannheim). Subsequent RP-HPLC analysis (5% acetonitrile in triethylammonium acetate 10 mM (pH 7) over 15 min.) was used to quantify the composition of the oligonucleotides (**1**; $\lambda_{max.}$= 271 nm, ε= 6000 $M^{-1}cm^{-1}$).

The accuracy of the PCR amplification was analyzed further by Saenger dideoxy-sequencing with Taq DNA polymerase[Innis, 1988 #1175] and chemical sequencing using the protocol of Maxam & Gilbert[Sambrook, 1989 #341] (see Standing Start Experiments and PCR amplification).

Polymerase screening, primer extension and scintillation counting

A series of commercially available thermostable polymerases was tested with the assay: Taq DNA polymerase and Tth DNA polymerase (Promega), Pfu DNA polymerase (cloned; Stratagene), Pwo DNA polymerase (Boehringer Mannheim), Vent™, Vent (exo-)™, Deep Vent™, Deep Vent (exo-)™, and 9°N DNA polymerase (all from New England Biolabs). All enzymes were incubated for 30 minutes at 72°C with a reaction mixture containing the supplier's reaction buffer, 1 μM of each dNTP (Promega), as well as either [^{3}H]-TTP (Amersham, 90-130 Ci/mmol) or [^{3}H]-dATP (ICN, 23 Ci/mmol), and 0.2 pmol of the primer/template complex (Figure 1A). The reaction was quenched with EDTA (0.5 M, pH 8.0), and the mixture was mixed with SPA™-beads (Amersham) or loaded into a Flashplate™ (NEN). After incubation at room temperature for 15-30 minutes, the samples were counted in a Topcount NXT Scintillation Microtiterplate Reader (Packard Instruments).

Standing Start Experiments and PCR amplification

The qualitative fidelity of Taq DNA polymerase was investigated by standing-start experiments and PCR amplification over 10 cycles. Primers with a 5'-^{32}P-label were obtained by enzymatic phosphorylation with polynucleotide kinase (New England Biolabs) and γ-^{32}P ATP (Amersham, >5000 Ci/mmol) and were ethanol-precipitated twice. After mixing primer (4 pmol), template (5 pmol) (Figure 3A), the corresponding dNTP (200 μM), and reaction buffer (supplied with

5

polymerase), the enzyme was added on ice and the sample cycled once (1 minute, 94°C; 30 seconds, 55°C; 10 minutes, 72°C). The reaction was quenched with EDTA and ethanol-precipitated before loaded on a 10% PAGE gel (7 M urea). The gel was analyzed with the MolecularImager® (BioRad).

For the PCR amplification experiments, primer (1 μM each, forward primer was 5'-^{32}P labeled) and template (20 nmol) (Figure 4A), as well as the appropriate dNTPs (200 μM each), were mixed with reaction buffer and Taq DNA polymerase (25 U) and adjusted to a final volume of 100 μl with water. The experiments were cycled (initial: 5 min., 94°C// 30 sec., 94°C; 30 sec., 55°C; 10 min., 72°C// polishing: 5 min., 72°C) ten times. During each annealing phase, an aliquot (8 μl) was taken, quenched with EDTA, and precipitated with ethanol. The samples were separated on a 10% PAGE gel (7 M urea) and analyzed with the MolecularImager® (BioRad).

The single-stranded oligonucleotides for DNA sequencing were produced by replacing the forward primer with a 5'-biotinylated analog. The quenched PCR reaction mixture was loaded onto a preequilibrated streptavidin-agarose column (Fluka Chemicals), washed with Tris-HCl (10 mM, pH 7.5) / 50 mM NaCl / 1 mM EDTA (pH 8.0), and the unbiotinylated strand was eluted with 0.2 M NaOH. After ethanol-precipitation and PAGE purification, the single-stranded material was used for Saenger dideoxy- and Maxam-Gilbert sequencing. Sequence analysis by the dideoxy method required some adjustment of the experimental conditions. When the temperature of the primer extension reaction was set at 70°C, the following dNTP/ddNTP/Mg concentrations were found suitable to sequence short oligonucleotides with Taq DNA polymerase: dA/ddA/Mg 15 μM / 2 mM / 2.24 mM; dT/ddT/Mg 15 μM / 3 mM / 3.24 mM; dG/ddG/Mg 15 μM / 0.5 mM / 0.74 mM; dC/ddC/Mg 15 μM / 1 mM / 1.24 mM. After annealing equal molar amounts of ^{32}P-labeled T7-primer with the purified template, all reaction mixtures were incubated with 5 pmol primer / template complex and 10 U of *Taq* DNA polymerase for 10 minutes. For the chemical sequence analysis, the protocol by Maxam-Gilbert for purine and pyrimidine analysis was employed. After 5'-^{32}P labeling of the amplified oligonucleotides with γ-^{32}P ATP (>5000 Ci/mmol) and T4 polynucleotide kinase (NEB), approximately 100,000 cpm (2 μl) were used per derivatization. Purine bases (A+G) were depurinated in 2.3 % formic acid (20-30 minutes at 37°C), followed by strand cleavage with piperidine (150 μl, 1 M, 90°C, 30 minutes). Pyrimidine bases were derivatized by hydrazine treatment (23 μl, 20°C, 4-10 minutes), quenched with hydrazine stop solution (200 μl) (Manatis

6

ref.) xxx and ethanol (750 µl); following lyophilization, the pyrimidine bases were cleaved with piperidine (100 µl, 1 M, 90°C, 30 minutes). All reaction mixtures from the dideoxy sequencing, as well as from the Maxam-Gilbert derivatization, were electrophorized in acrylamide (10%, 7 M urea) and analyzed by autoradiology.

Gel Fidelity Assay for (1)

Once a polymerase that accepted incorporated **1** was identified by the high-throughput screen, kinetic data for the incorporation of **1** by the polymerase were determined by conventional single nucleotide incorporation[Boosalis, 1987 #1221]. The same primer/template complexes (Figure 3A) as in the previous qualitative analysis were used. In the absence of dNTP, primer (^{32}P-labeled; 20 pmol), template (25 pmol), and Taq DNA polymerase (0.1 U) were mixed in reaction buffer, denatured at 94°C (1 min.), and annealed at 55°C (60 min.). The experiment was initiated upon addition of the appropriate dNTP at 55°C and the reaction was quenched after a previously determined time to ensure a maximum of 20% of the primer was extended (single turnover conditions). After separation by PAGE, the products were quantified using the MolecularImager® and MultiAnalyst® software (BioRad). Kinetic parameters were determined by plotting [dNTP] versus [dNTP]/v (v is the primer extension velocity (ratio (n+1)/(n+(n+1)) per second) (Hanes-Woolf plot). *what is n xxx?*

Results / Discussion:

In vitro screening assay / SPA:

The scintillation proximity assay (SPA™, AmershamPharmacia Biotech, Arlington Height, IL)[Cole, 1996 #1208] detects the incorporation of tritium-labelled dNMP into a product that is attached to a streptavidin-coated plastic beads, which contain a scintillant. Because of the relatively low energy of the tritium decay particle, radiolabel is counted more efficiently when it decays in the vicinity of the scintillant. The primer is 5'-biotinylated, and binds with the template to the bead with high affinity. Unincorporated [^{3}H]-dNTP in solution reflects background, usually in the range of 1-5% of the positive control, while [^{3}H]-dNTP added to the primer is counted efficiently. An analogous strategy can be implemented using Flashplates (NEN™ Life Science Products, Boston, MA), a

7

microtiter plate where the streptavidin is on the surface of the plastic wells and scintillant is embedded in the plastic. Setting up the primer extension reactions requires ca. 15 minutes and incubation with polymerase requires 30-60 minutes, making the plate screening 96 variants available for counting (typically 4 hours) in less than two hours.

To test the screen, we looked for polymerases that could handle pseudothymidine in the template (**1**). The C-nucleoside analog of thymidine (**2**) was prepared by organic synthesis from the commercially available pseudouridine (Methods and Materials). Oligonucleotides containing **1** were prepared by phosphoramidite-based DNA synthesis. A template (30-mer; Figure 1A) was designed that contains no 2'-deoxyadenosine between the 3'-end of the primer (position 21) and **1** (X; position 25). The template was annealed with a 5'-biotinylated primer (P/T 21/30) and incubated with ^{3}H-TTP. Tritium becomes attached to the bead only if dATP is successfully incorporated opposite **1**, permitting further elongation of the primer beyond position 25 and the incorporation of tritium-labeled TTP opposite position 27. As a control, analogous primer extension reactions were run in the absence of various dNTPs to learn whether the polymerase enabled tritium incorporation via mismatches (Figure 1B). Thus, experiment #1 tests for primer extension beyond the unnatural nucleotide, experiment #2 tests the specificity of the incorporation in the absence of dATP (the appropriate hydrogen-bonding partner for **1**), while experiment #3 measures the background radioactivity of the reaction mixture.

The incorporation of the correct nucleoside triphosphate opposite **1** can also be measured by using tritium-labeled dATP in a second, independent set of reactions (Figure 1B #4/5). The question of sensitivity and reproducibility of the method was addressed by a series of experiments with serial dilutions of a sample of 9°N DNA polymerase (xxx need a reference for this polymerase). Measurements of counts per minute for a given reaction was reproducible within 10%.

As a proof of concept, a set of commercially available thermostable polymerases suitable for the polymerase chain reaction were then screened for their ability to incorporate **1** using the P/T 21/30 complex. A set of five reactions was run for each polymerase. After quenching with EDTA, the amount of tritium incorporation was determined by scintillation counting on SPA beads (Figure 1C).

Interestingly, misincorporation correlated directly with the exonuclease activity of the polymerase. Proofreading is usually viewed as a mechanism to ensure fidelity of DNA replication;

8

with the unusual nucleotide, proofreading may offer a mechanism for lowering the fidelity for Vent, Deep Vent, Pfu, and the Pwo DNA polymerases. Based on previous observations that polymerases pause after formation of base pairs involving unnatural nucleotides, we hypothesize that the halt allows the exonuclease activity (k_{-1}, Figure 2) to become kinetically competent, resulting in removal of the nucleotide at position 25 from the primer. Upon repetition of the polymerization step (k_{+1}) with a fresh nucleoside triphosphate, the pausing will repeat since the causative factor, the unnatural nucleotide, is located in the template. Eventually, the stalled complex either disassociates (k_{off}) or is extended beyond the "misincorporation site" (k_{+2}). However, the energetic advantage of the correct hydrogen-bonding partner, contributing to the specificity of correct nucleotide incorporation opposite **1** in the template becomes statistically insignificant, resulting in increased misincorporation. Conversely, polymerases with reduced or no exonuclease activity (Taq, Tth, 9°N, Vent (exo⁻), Deep Vent (exo⁻)) will favor the incorporation of the correct nucleoside triphosphate, based on the absence of destabilization of the newly forming base pair resulting from incorrect hydrogen bonding pattern. The significance of the exonuclease activity is particularly apparent when comparing the primer extension of Vent and Deep Vent DNA polymerase with their (exo⁻)-mutants.

Standing start experiments:

Based on the screening using the scintillation proximity assay, the DNA polymerases from *Thermus aquaticus* and *Thermus thermophilus* were chosen for further standing start experiments [Boosalis, 1987 #1221]. Two primer/template complexes (Figure 3A; P/T 24/30), carrying either **1** or deoxyadenosine at position 25 were prepared. An extended reaction time of 30 minutes allowed the polymerase to undergo several cycles of polymerization (multiple hit conditions). The resulting gel, while not suitable for quantitative analysis, identified nucleoside triphosphates that might be misincorporated opposite **1** in the template, as well as the competitors of ψTTP incorporation opposite dA in the template. *Tth* DNA polymerase performed rather poorly, showing significant misincorporation under these conditions (data not shown). *Taq* DNA polymerase (Figure 3B), however, extended the primer by placing a purine nucleotide, preferentially deoxyadenosine, opposite to **1** in the template. *Taq* polymerase also incorporated ψTTP efficiently opposite dA in the template. The only nucleoside triphosphate misincorporated to a significant extent was dATP (**xxx Stefan;**

9

did I get this right?). The extended primer was most likely the result of non-template directed polymerization by Taq DNA polymerase, creating a 3'-A overhang.

Steady-state kinetics of ψT-incorporation by gel fidelity assay

A quantitative comparison of the efficiency of incorporation of ψT and its natural competitors was made by measuring the single-nucleotide insertion by varying the dNTP concentration under single-turnover conditions. Applying the same primer-template complex as above (P/T 24/30), a series of kinetic studies was carried out using steady-state methods [Goodman, 1993 #952][Creighton, 1995 #1369]. Based on the above qualitative findings, kinetic data were determined only for the possible competitors and **1** itself. The insertion of dATP and dGTP opposite ψT in the template and the incorporation of TTP, ψTTP, and dATP opposite dA were studied. Kinetic parameters derived from the data are presented in Table 1. The insertion of dATP opposite ψT in the template is 50-fold more efficient than insertion of dGTP. This discrimination is smaller than that observed with standard nucleotides, which show discrimination ratios between 1×10^3 and 1×10^4. In the opposite direction, incorporation of ψTTP opposite dA in the template is only 2-fold less efficient than TTP, the natural hydrogen bonding partner of dA. The insertion of dATP, on the other hand, is 3×10^2-fold lower and, as suggested previously, probably the result of non-template specific polymerization.

PCR amplification

Once performance with good fidelity and efficiency under single nucleotide incorporation conditions was achieved, primer elongation by *Taq* DNA polymerase under PCR-like conditions was tested. A reaction mixture containing a standard oligonucleotide (51-mer; Figure 4A) flanked by the T3- and the T7-primer binding sequence was prepared and PCR amplified in the presence of dATP, dCTP, dGTP, and ψTTP. Two control reactions, one substituting ψTTP with standard TTP, the other performed in the absence of any thymidine analog, were run simultaneously. A sample was taken after each cycle and analyzed by PAGE (Figure 4B). The reaction mixture lacking thymidine or its analog generated no full-length product. The primer extension reaction was terminated completely at the position of the first adenosine in the template. In the presence of ψTTP, the desired full-length

10

oligonucleotide represented the primary product. Although up to three ψT's in succession have to be incorporated in order to produce full-length product, only a small fraction of the extended primer appeared as an early termination product. Further PCR amplification up to 10 cycles gave similar results (data not shown) and suggested that Taq DNA polymerase could be used for PCR-like amplification of sequences containing ψT.

Amplification fidelity / Sequencing

In addition to the evidence of ψT-incorporation by *Taq* DNA polymerase under conditions of PCR amplification, further proof for sequence-specific amplification of an oligonucleotide in the presence of ψTTP was sought. Modifying the original PCR protocol, the T7 primer was replaced by its 5'-biotinylated analog. After the regular PCR cycles, the amplified material was loaded onto a streptavidin column and washed. Single-stranded oligonucleotides were isolated by elution with sodium hydroxide solution and purified from the excess primer by PAGE. This material was used directly for Saenger dideoxy-sequencing and Maxam-Gilbert sequencing, as well as in enzymatic digestion. A complete digestion of the PCR product into the nucleosides by phosphodiesterase and alkaline phosphatase followed by quantification by reverse phase HPLC analysis gave the predicted ratio of nucleosides. The dideoxy-sequencing of the PCR product confirmed sequence identity with the original template (Figure 5A). Excluding the first few positions after the primer which could not be read conclusively due to high background, the sequence pattern was consistent with the original template. The purine/pyrimidine distinction, seen in Maxam-Gilbert sequencing (Figure 5B), also indicated identical sequences after 10 cycles of PCR amplification. The positions of ψT were identified easily by the missing bands in the C+T pattern. Clearly, the C-nucleoside does not undergo strand cleavage upon hydrazine/piperidine treatment and therefore does not appear as a distinct band in the sequencing gel.

Conclusions

In the present paper, we have successfully tested a new *in vitro* screening assay for thermostable DNA polymerases. The combination of a set of primer extension reactions and the SPA technology for its rapid and simple quantification represents a novel method to search for polymerases

11

that can utilize functionalized or structurally modified nucleotides. The method is applicable to any nucleoside analog, whether integrated in an oligonucleotide or present as the triphosphate. To further simplify and speed up the procedure, the entire assay has been transferred into 96-well microtiter plates (Flashplates), making it suitable for high-throughput screening of protein libraries.

The best polymerase obtained from the screen was *Taq* DNA polymerase. Conventional kinetic studies were then performed to demonstrate that the polymerase does perform with **1** as expected from the results of the screen. The fidelity of synthesis of the A-ΨT base pair is lower than with the conventional base pairs. It is sufficient, however, for the amplification of typical oligonucleotides for *in vitro* selection (up to 150 nucleotides over 10 cycles).

The results from the kinetic experiments can be put into perspective with previously published data on related nucleoside analogs such as difluorotoluene deoxynucleoside[Moran, 1997 #757][Moran, 1997 #992]. While the structural distortion of the nucleoside analog as seen for **1** reduces the accuracy of replication approximately two-fold as measured by kinetic experiments (Table 1, dA template vs. TTP and ψTTP), removal in addition of the Watson-Crick hydrogen bonding pattern in the difluoro-compound causes the accuracy for incorporation of dFTP compared to TTP opposite dA drop by a factor of 40[Moran, 1997 #992]. Although quantitative comparsion would require the experiments to be measured in parallel, these preliminary data may help addressing the question of the individual contribution of hydrogen bonding and geometry to the overall fidelity of DNA polymerases.

SIMPLE ONE POT SYNTHESIS OF A
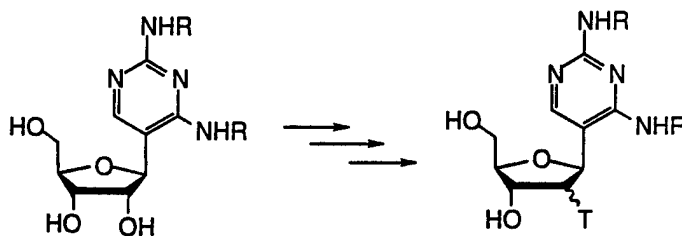# 2'-TRITIUM LABELED C-DEOXYNUCLEOSIDE

Stefan Lutz and Steven A. Benner*
Department of Chemistry, University of Florida, FL, 32611, USA

The deoxygenation and 2'-labeling of a C-ribonucleoside
by reductive elimination with tri-*n*-butyltin hydride[^{3}H] in
a one-pot reaction is described.

Simple One Pot Synthesis of a
2'-Tritium Labeled C-Deoxynucleoside

Stefan Lutz and Steven A.Benner*

Department of Chemistry, University of Florida, Gainesville FL, 32611, USA

Abstract: The deoxygenation and 2'-labeling of a C-ribonucleoside by reductive elimination with tri-*n*-butyltin hydride[^{3}H] in a one-pot reaction is described. The approach is a safe, simple, efficient, and general method for 2'-labeling of nucleosides.

Recently, work in many laboratories has shown that a wide range of heterocyclic bases can contribute to molecular recognition between two strands of DNA and RNA [1-7]. Nucleoside analogs bearing these heterocycles and labeled with radioactive isotopes would be valuable tools to expand our understanding of this range. With natural nucleosides, a variety of enzymatic [8] and chemical [9,10] procedures have been reported for introducing isotopes of hydrogen, including deuterium [11-15] and tritium. These procedures are often inconvenient, requiring multiple steps, giving low yields, causing anomerization, or generating tritium gas as a side product.
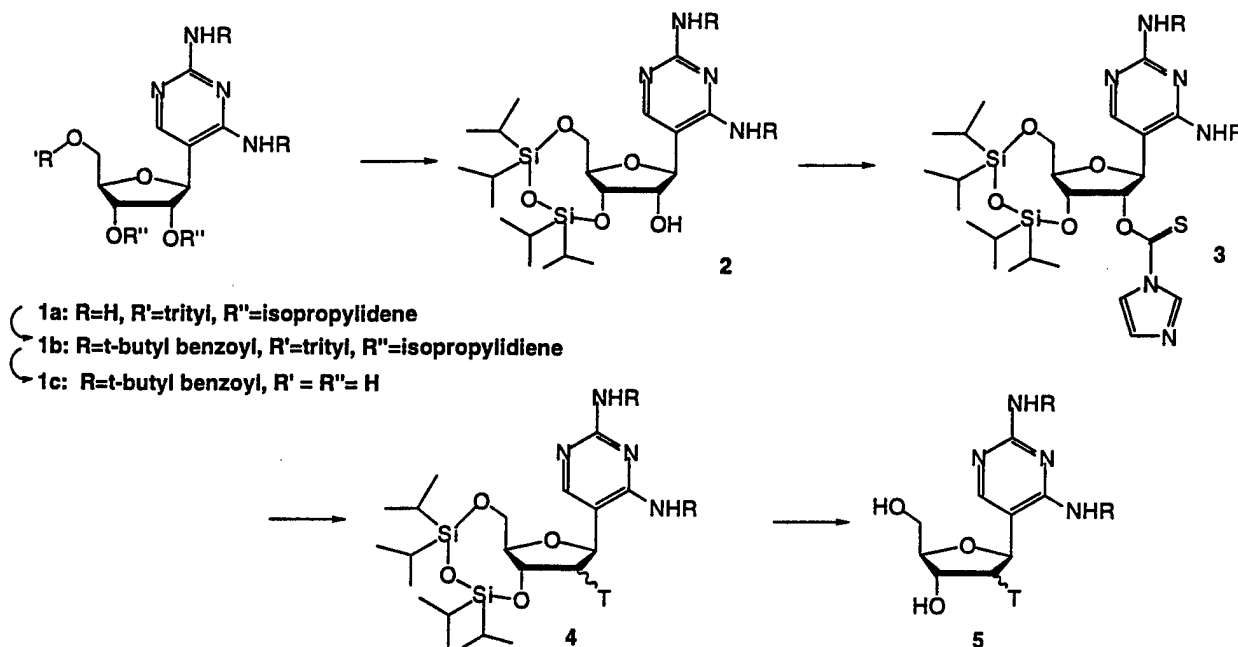

1a: R=H, R'=trityl, R"=isopropylidene
1b: R=t-butyl benzoyl, R'=trityl, R"=isopropylidiene
1c: R=t-butyl benzoyl, R' = R"= H

Figure 1: The synthetic scheme for a [2'-^{3}H]-2'-deoxy-C-nucleoside from its ribonucleoside is shown

We report here a procedure that exploits the known ability of deuterated tri-*n*-butyltin hydride to transfer a hydrogen specifically to a radical at the C-2' position of a nucleoside[15,16]. The procedure was adapted first to incorporate a tritium from tri-*n*-butyltin tritiide (Moravek Biochemicals Inc., 577 Mercury Lane, Brea, CA 92821 USA), which was prepared by reacting tri-*n*-butyltin hydride with MeMgBr, and hydrolyzing the intermediate with tritiated water[17,18]. Distillation yielded tri-*n*-butyltin hydride (>96% pure as determined by HPLC), which was stored neat at -20°C.

To explore the scope of this reaction with some of the more novel nucleoside analogs, we developed a one pot procedure applying the reagent to 2,4-diamino-5-(β-D-ribofuranosyl)-pyrimidine (**1a**)[19-21], a non-standard pyrimidine that presents a hydrogen bond donor-acceptor-donor pattern when forming a Watson-Crick base pair. The exocyclic amino groups of **1a** were protected as amides via DMAP-catalyzed reaction with p-*tert*butyl benzoyl chloride (pyridine, 0°C, 2 h). The isopropylidene- and trityl protection groups were removed (10% HCl in MeOH, 2 hr, RT), and the ribose derivative **1c** was silylated at the 3'- and 5'-positions with dichloro-tetraisopropyl disiloxane (pyridine, 4 h, RT) to yield **2**. Reaction of the 2'-hydroxyl group with 2,2'-thiocarbonyl-diimidazole in DMF yielded the thiocarbonylimidazolide derivative **3**.

In the presence of 2,2'-azo-bisisobutyronitrile and toluene, **3** was refluxed with the radiolabeled tri-*n*-butyltin hydride for 1 hour, yielding **4**. Treatment with tetrabutylammonium fluoride (1 M, THF) gave the 2'-[^{3}H]-2'-deoxy-nucleoside **5**, which was purified by silica-gel chromatography (8% MeOH/CHCl$_3$). In preliminary studies, all intermediates were isolated and characterized[22]. In its one pot version, the yield of **5** was 75% (based on **1c**); the radioactive yield based on the specific activity of the [^{3}H]-tri-*n*-butyltin hydride was 50%. Radiolabeled product was analyzed by thin layer chromatography and was identical to the reference material. In addition, the TLC plates were exposed to a tritium-sensitive imager screen. Radioactivity was detected only in the product spot.

This efficient one pot synthesis of 2'-radioisotopic labeled deoxynucleoside is expected to be applicable to a wide range of nucleoside analogs, given the known tolerance of tri-n-butyltin hydride deoxygenation to a variety of heterocycles. These compounds can be readily converted into the corresponding triphosphates by chemical[23] and enzymatic methods[24,25]. Applying this strategy of 2'-radiolabeling to a variety of nucleosides or nucleosides analogs makes these compounds valuable tools for high throughput screening techniques such as the Scintillation Proximity Assay (SPA™, Amersham Life Science Inc.)[26].

REFERENCES AND NOTES

1. Moran, S.; Ren, R. X. F.; Rumney, S.; Kool, E. T. *J.Amer.Chem.Soc* **1997**, *119*, 2056-2057.
2. Moran, S.; Ren, R. X. F.; Kool, E. T. *Proc.Natl.Acad.Sci.USA* **1997**, *94*, 10506-10511.
3. Roberts, C.; Bandaru, R.; Switzer, C. *J.Amer.Chem.Soc* **1997**, *119*, 4640-4649.
4. Tarasow, T. M.; Tarasow, S. L.; Eaton, B. E. *Nature* **1997**, *389*, 54-57.
5. Seela, F; Debelak, H; Usman, N; Burgin, A; Beigelmann, L. *Nucl.Acids Res.* **1998**, *26*, 1010-1018.
6. Seela, F.; Wei, C. *Helv.Chim.Acta* **1997**, *80*, 73-85.
7. Heidenreich, O.; Benseler, F.; Fahrenholz, A.; Eckstein, F. *J.Biol.Chem.* **1994**, *269*, 2131-2138.
8. Chen, A. K.; Bhan, A.; Hopper, S.; Abrams, R.; Franzen, J. S. *Biochemistry* **1974**, *13*, 654-661.
9. David, S.; Eustache, J.; Rouzeau, C. *C.R. Acad. Sci.Ser.C* **1970**, *270*, 1821-1824.
10. David, S.; Auge, C. *Carbo.Res.* **1973**, *28*, 125-128.
11. Fraser-Reid, B.; Radatus, B. *J.Amer.Chem.Soc* **1971**, *93*, 6342-6344.
12. Pathak, T.; Bazin, H.; Chattopadhyaya, J. *Tetrahedron* **1986**, *42*, 5427-5441.
13. Cook, G. P.; Greenberg, M. M. *J.Org.Chem.* **1994**, *59*, 4704-4706.
14. De Voss, J. J.; Hangeland, J. J.; Townsend, C. A. *J.Org.Chem.* **1994**, *59*, 2715-2723.
15. Robins, M. J.; Wilson, J. S.; Hansske, F. *J.Amer.Chem.Soc* **1983**, *105*, 4059-4065.
16. Barton, D. H. R.; Subramanian, R. *J.Chem.Soc.Chem.Commun.* **1976**, *21*, 867-868.
17. Lee, S. Y. *Bull.Singapore Natl. Inst. Chem.* **1990**, *18*, 115-120.
18. Albert, H. J.; Neumann, W. P. *Synthesis - Stuttgart* **1980**, 942-943.
19. Chu, C. K.; Reichmann, U.; Watanabe, K. A.; Fox, J. J. *J.Org.Chem.* **1977**, *42*, 711-714.
20. Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. *Nature* **1990**, *343*, 33-37.
21. Piccirilli, J. A.; Krauch, T.; MacPherson, L. J.; Benner, S. A. *Helv.Chim.Acta* **1991**, *74*, 397-406.
22. Protons on the ribose were assigned based on nOe experiments and homonuclear COSY. The stereochemistry at the anomeric center was determined by nOe diff. experiments.
   Compound **1b**: ^{1}H-NMR (500MHz, CDCl$_3$) δ: 11.8 (s, 1H, NH); 10.2 (s, 1H, NH); 8.52 (s, 1H, H-6); 8.08 (d, 2H, arom.H); 7.85 (d, 2H, arom.H); 7.51 (d, 2H, arom.H); 7.49 (d, 6H, arom.H); 7.44 (d, 2H, arom.H); 7.29 (t, 6H, arom.H); 7.22 (q, 3H, arom.H); 5.58 (d, 1H, H-1', J=3.3 Hz.); 4.87 (q, 1H, H-2', J= 3.3 / 6.3 Hz.); 4.57 (t, 1H, H-3', J= 6.3 / 5.6 Hz.); 4.29 (q, 1H, H-4', J= 5.6 / 2.5Hz.); 3.37 (m, 2H, H-5'$_{B/A}$); 1.70 (s, 3H, isoprop.CH$_3$); 1.38 (s, 3H, isoprop.CH$_3$); 1.34 (s, 9H, t-bu); 1.32 (s, 9H, t-bu) ppm ; ms: m/z 845 (M$^+$ +1), 243 (M$^+$ -601), 161 (M$^+$ -683)
   Compound **1c**: ^{1}H-NMR (300MHz, d-MeOH) δ: 8.67 (s, 1H, H-6); 7.92 (dd, 4H, arom.H); 7.52 (q, 4H, arom.H); 4.91 (m, 1H, H-1); 4.12 (m, 3H, H-2', H-3', H-4'); 3.79 (dd, 1H, H-5'$_B$, J= 2.7 / 12.2 Hz); 3.70 (dd, 1H, H-5'$_A$, J= 4.2 / 12.2 Hz); 1.39 (s, 9H, t-bu); 1.32 (s, 9H, t-bu); ppm ; ^{13}C-NMR (300MHz, d-MeOH) δ: 167.1; 167.0; 158.0; 157.9; 157.8; 157.4; 157.1; 132.2; 132.2; 129.0; 128.9; 126.8; 126.7; 118.1; 88.4; 79.3; 77.6; 73.2; 63.3; 35.9; 31.5 ppm

Compound **2**: ^{1}H-NMR (500MHz, CDCl$_3$) δ: 9.95 (s, br, 1H, NH); 9.00 (s, 1H, NH); 8.71 (s, 1H, H-6); 7.88 (dd, 4H, arom.H); 7.47 (d, 4H, arom.H); 4.84 (d, 1H, H-1'; J=5.5 Hz); 4.35 (t, 1H, H-3', J=6.5 Hz); 4.16 (m, 1H, H-2'); 4.14 (dd, 1H, H-5'$_B$, J= 3.2 / 11.6 Hz); 3.99 (hex, 1H, H-4', J= 3.2 / 7.0 / 6.5 Hz); 3.93 (q, 1H, H-5'$_A$, J= 7.0 / 11.6 Hz); 1.34 (s, 18H, t-bu); 1.11-0.96 (m, 28H, isoprop.) ppm ;

Compound **3**: ^{1}H-NMR (300MHz, CDCl$_3$) δ: 8.56 (s, br, 1H, NH); 8.32 (s, 1H, NH); 8.12 (s, 1H, H-6); 7.90 (m, 4H, arom. H); 7.69 (s, 1H, imid.); .7.60 (s, 1H, imid.); 7.48 (m, 4H, arom.H); 7.03 (s, 1H, imid.); 6.14 (d, 1H, H-2'; J= 5.0 Hz); 5.55 (s, 1H, H-1'); 4.45 (q, 1H, H-3'; J= 5.0 / 8.7 Hz); 4.18 (m, 1H, H-5'$_B$, J= 12.3 Hz); 4.02 (m, 2H, H-4' / H-5'$_A$); 1.31 (s, 9H, t-bu); 1.25 (s, 9H, t-bu); 1.06-0.8 (m, 28H, isopropyl); ppm;

Compound **4**: ^{1}H-NMR (300MHz, CDCl$_3$) δ: 10.0 (s, br, 1H, NH); 8.90 (s, br, 1H, NH); 8.44 (s, 1H, H-6); 7.84 (dd, 4H, arom.H); 7.47 (t, 4H, arom.H); 5.17 (q, 1H, H-1', J= 8.4 / 6.3 Hz); 4.50 (m ,1H, H-3'); 4.13 (dd, 1H, H-5'$_B$, J= 3.6 / 11.5 Hz); 3.97 (m, 1H, H-4'); 3.75 (m, 1H, H-5'$_A$, J= 8.7 / 11.5 Hz); 2.39 (m, 1H, H-2'$_B$); 2.30 (m, 1H, H-2'$_A$); 1.33 (s, 9H, t-bu); 1.32 (s, 9H, t-bu); 1.10-0.97 (m, 28H, isopropyl); ppm; ms: m/z 811 (M$^+$ +23), 789 (M$^+$ +1), 161 (M$^+$ -627)

Compound **5**: ^{1}H-NMR (300MHz, d-MeOH) δ: 8.48 (s, 1H, H-6); 7.95 (dd, 4H, arom.H); 7.57 (dd, 4H, arom.H); 5.29 (q, 1H, H-1', J= 5.1 / 10.5 Hz); 4.32 (m, 1H, H-3'); 4.06 (m, 1H, H-4'); 3.73 (dd, 1H, H-5'$_B$, J= 3.9 / 12.0 Hz); 3.64 (dd, 1H, H-5'$_A$, J= 5.4 / 12.0 Hz); 2.38 (dd, 1H, H-2'$_B$, J= 5.1 / 13.0 Hz); 2.20 (hex, 1H, H-2'$_A$, J= 6.0 / 13.0 / 10.5 Hz); 1.36 (s, 18H, t-bu); ppm ; ^{13}C-NMR (300MHz, d-MeOH) δ: 167.2; 167.1; 158.5; 158.2; 158.0; 157.7; 157.5; 132.4; 132.0; 128.9; 126.9; 126.7; 119.4; 89.9; 77.1; 73.5; 63.6; 54.2; 41.8; 36.0; 35.9; 31.5; ppm. ms: m/z 569 (M$^+$+23), 547 (M$^+$ +1), 161 (M$^+$ -385).

23. Ludwig, J.; Eckstein, F. *J.Org.Chem.* **1989**, *54*, 631-635.

24. Lieberman, I.; Kornberg, A.; Simms, E. S. *J.Biol.Chem.* **1955**, *215*, 429-440.

25. Skoblov, Y. S.; Korolev, A. E.; Maslova, R. N. *Russ.Chem.Rev.* **1995**, *64*, 799-807.

26. Cole, J. L. *Methods Enzym.* **1996**, *275*, 310-328.